Achieving Successful VLC Signal Reception Using A Rolling Shutter Image Sensor While Driving at 40 km/h

Shunki Kamiya, *Graduate Student Member, IEEE*, Zhengqiang Tang, *Graduate Student Member, IEEE*, Takaya Yamazato, *Senior Member, IEEE*, Masayuki Kinoshita, *Member, IEEE*, Koji Kamakura, *Member, IEEE*, Shintaro Arai, *Member, IEEE*, Tomohiro Yendo, *Member, IEEE*, and Toshiaki Fujii, *Member, IEEE*

***Abstract*—This paper presents the successful reception of visible light communication (VLC) signals, transmitted from an LED array, by an image sensor that utilizes a rolling shutter while driving at 40 km/h. The rolling shutter image sensors used in commercial cameras, such as those found in smartphones and dashcams, capture images line by line at a fast rate, allowing for VLC signal reception in vehicular environments. By analyzing the relationship between the signal reception rate for each line and the frame rate, we demonstrate that parallel VLC signals transmitted from the LED array can be received even while in motion. Notably, to our knowledge, this is the first time that an automobile moving at 40 km/h has successfully received a VLC signal.**

***Index Terms*—Visible light communication, intelligent transport systems, rolling shutter image sensor, LED array.**

I. Introduction

VEHICLE-TO-EVERYTHING (V2X) visible light communication (VLC) technology [1], [2] improves the efficiency and safety of the traffic system. In V2X VLC, data is transmitted by LED light sources such as LED traffic lights, the head and tail lights of vehicles, digital signage, and so on [3], [4], [5], [6]. One of the major advantages of V2X VLC is its ease of implementation because these light sources are already installed on the road. Therefore, transmitting traffic information in this way is expected to be a useful communication method to assist automated driving.

This paper explores the use of rolling shutter (RS) image sensors, which are inexpensive and widely available in smartphones, for V2X VLC signal reception. Previous studies on V2X VLC often used global shutter (GS) image sensors as receivers due to their high acquisition speed, as noted in [7]. We clarify RS and GS image sensors with respect to their exposure mechanism. Fig. 1 shows the exposure mechanism of RS and GS image sensors. A row of pixels in an image sensor is considered to be a scan line. Based on the resolution of the image sensor, the number of scan lines in an image sensor varies from hundreds to thousands. Both RS and GS image sensors are exposed in scan line units. The biggest difference between these two sensors is that the RS image sensor sequentially exposes each scan line from the top to the bottom of the sensor. A certain time interval exists between the exposure start times of adjacent scan lines in RS image sensors, as shown in Fig. 1. On the other hand, the GS image sensor exposes all scan lines at the same time. For GS image sensors, the start and end time of the exposure for each scan line is the same. Therefore, there is no time difference between line-to-line exposures in GS image sensors. However, GS high-speed image sensors are expensive and not commonly used. On the other hand, VLC using RS image sensors has been widely studied in static environments [8], [9], [10], [11], [12], [13]. In [8], the RS effect [14] is used to improve the data rate of VLC. In [9], a monitor (e.g., advertisement display) is used as a transmitter to achieve a display rate tens to hundreds of times higher than the frame rate, achieving a data rate of 1 kbps. In addition, systems like LinkRay have actually been implemented in public places [15], [16].

Here, we describe the background of V2X VLC technology and its applications in vehicular environments. As a promising

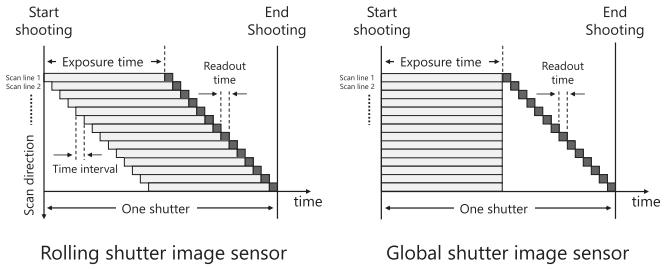


Fig. 1. The exposure mechanism of RS and GS image sensors.

Manuscript received 17 March 2023; revised 16 May 2023; accepted 12 June 2023. Date of publication 19 June 2023; date of current version 28 June 2023. This work was supported by the Okawa Foundation of Information and Telecommunications. *(Corresponding author: Zhengqiang Tang.)*

Shunki Kamiya, Zhengqiang Tang, Takaya Yamazato, and Toshiaki Fujii are with the Nagoya University, Nagoya-shi 464-8601, Japan (e-mail: skamiy@katayama.nuee.nagoya-u.ac.jp; tang@katayama.nuee.nagoya-u.ac.jp; yamazato@ieee.org).

Masayuki Kinoshita and Koji Kamakura are with the Chiba Institute of Technology, Narashino-shi 275-8588, Japan.

Shintaro Arai is with the Okayama University of Science, Okayama-shi 700-0005, Japan.

Tomohiro Yendo is with the Nagaoka University of Technology, Nagaoka-shi 940-2188, Japan.

Digital Object Identifier 10.1109/JPHOT.2023.3287211

application of VLC, V2X VLC technology attracts considerable attention. Vehicle VLC technology involves the use of high-intensity LED lights, such as headlights, brake lights, or traffic lights, to transmit data. The data are generally encoded by modulating the intensity or the flash frequency of the lights. To receive the transmitting data, an image sensor (camera) or a photodiode is usually used to detect and recover the light signals. In vehicular environments, VLC technology can be used for a variety of applications, including V2V and V2I (vehicle-to-infrastructure) communication, autonomous driving, and advanced driver assistance systems (ADAS) [17]. For example, VLC can enable communication between vehicles to facilitate cooperative driving, such as platooning [18], where multiple vehicles follow each other closely to reduce air resistance and improve fuel efficiency [19]. VLC can also be used to communicate with infrastructure, such as traffic lights, to improve traffic flow and reduce congestion [20]. Therefore, we consider that vehicle VLC technology is a promising wireless communication technology that has the potential to revolutionize vehicular environments. There are some other studies on the use of RS image sensors for VLC in the driving environment. In [21], an RS image sensor is used to estimate the position of a moving vehicle. In [22], a filter is used in the transmitter to enable communication over long distances and with multiple users. In addition, some studies examined the use of LEDs and RS image sensors for VLC [23], [24], [25]. However, most of the research in moving environments using RS image sensors focuses on vehicle position estimation and noise reduction.

This study aims to realize VLC between vehicles traveling at speeds of up to 40 km/h, which is the normal speed for urban roads, using an LED array as the transmitter and an RS image sensor as the receiver. We address two issues that arise due to the characteristics of the RS image sensor. First, the camera's shooting speed is not fast enough to capture the high-speed blinking LEDs. Second, the positional relationship between the transmitter and receiver varies depending on the moving environment. To overcome these problems, we propose a V2X VLC scheme that demodulates data in a vehicular environment while taking into account the characteristics of the RS image sensor.

The rest of the paper is organized as follows. In Section II, we describe the proposed VLC system model. In Section III, we explain the configuration of the transmission signal, which takes into account the characteristics of the image sensor. In Section IV, we describe the characteristics of the received image obtained when the transmitted signal is captured by the rolling shutter image sensor and how it is captured. In Section VI, we detail the characteristics of the received image and the demodulation process. Finally, in Section VII, we provide a conclusion to this paper.

II. Proposed Vehicular VLC System

The proposed system model is shown in Fig. 2. In this study, a 16×16 LED array is used to transmit optical signals. The input transmission data is added with a pilot sequence. The added pilot sequence is used to detect the beginning of the data transmission.

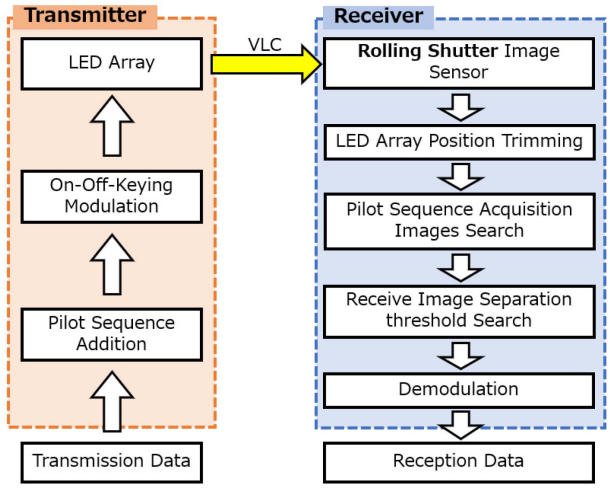


Fig. 2. System model.

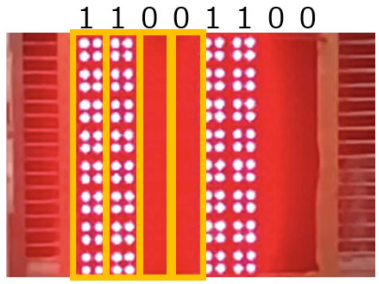


Fig. 3. Transmission frame configuration.

We use ON–OFF keying (OOK) as the modulation method. The binary data '1' and '0' are indicated by the ON and OFF of the LED, respectively. The LED array blinks LED patterns at a frequency of 500 Hz to transfer the optical signal.

The transmitted optical signals travel through the spatial channel and are captured by the receiver camera of the vehicle. The receiver captures the blinking patterns of the LED array as images using an RS image sensor. Then, the receiver trims out the LED array portion from the captured image. Next, we identify the pilot sequence images to find the data images. After that, the trimmed LED array is divided into segments in a range representing 1 b. Finally, we extract the pixel values of the divided segments and recover data by threshold determination.

III. Transmission Signal Configuration

This section describes the composition of the transmission signal. In this study, the communication between the transmitter and the receiver is asynchronous. This means the start of transmission does not necessarily coincide with the start of the camera capturing.

A. Transmission Frame Configuration

In this study, we use a 16×16 array of LEDs as the transmitter, as shown in Fig. 3. The transmitted data is blocked as

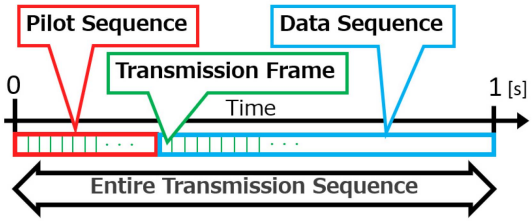


Fig. 4. Composition of entire transmission sequence, including pilot sequence and data sequence.

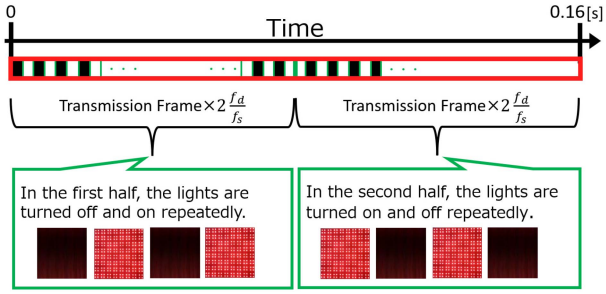


Fig. 5. Pilot sequence configuration. ON-OFF pattern switches between the first and second half.

a series of binary symbols, each of which contains N bits in a transmission frame. The i-th data block is defined as $\boldsymbol{d}(i)$ as follows.

$$\boldsymbol{d}(i) = \{d_0(i), \cdots d_{N-1}(i)\},\ N = 8, \tag{1}$$

where $d_k(i)$ represents 1 b of data, and $d_k(i)$ is expressed as follows.

$$d_k(i) \in \{0, 1\} \quad (k = 0, 1, \cdots N-1). \tag{2}$$

Here, the parameter k denotes the k-th bit of data in a transmission frame.

The data is modulated by OOK and is transmitted by blinking LEDs. We show a transmission frame displaying a data block $\boldsymbol{d}(i) = [1, 1, 0, 0, 1, 1, 0, 0]$ in Fig. 3. To receive the signal from a single transmission frame, we vertically divided the LED array using 16 × 2 LEDs to transmit 1 b of data. In this case, the receiver RS image sensor is able to horizontally capture some lines of the transmission frame.

B. Sequence Configuration

The configuration of the transmission sequence is shown in Fig. 4. The transmission sequence consists of a pilot sequence and a data sequence. As mentioned above, we use the pilot sequence to determine the start of the data sequence. In the pilot sequence, all LEDs alternately switch between ON and OFF states. The detail of the pilot sequence configuration is shown in Fig. 5. We divide the pilot sequence into the first half and the second half. In the first half, we send transmission frames repeating all lights OFF and all lights ON. In the second half, we send the opposite patterns of the first half which repeats all lights ON and all lights OFF. Let f_s [fps] denote the shooting speed, f_d [Hz] denote the LED display speed, and M_p denote

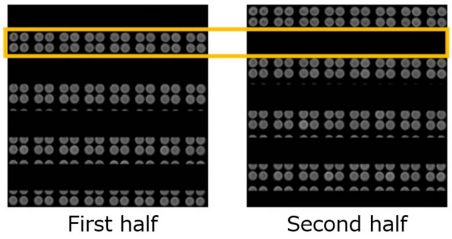


Fig. 6. Two different images taken of the pilot sequence. If a bright line is captured in the first half, a dark line is captured in the second half.

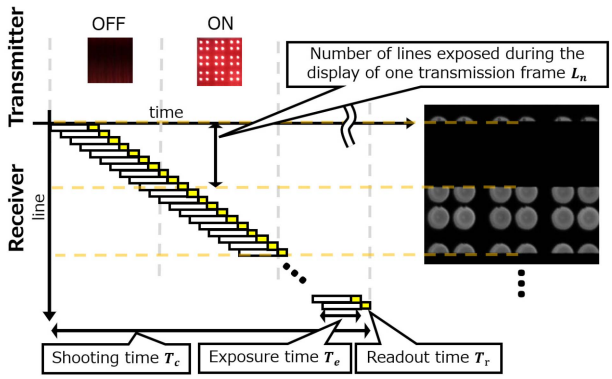


Fig. 7. Timing chart for photographing a high-speed blinking LED array using a rolling shutter image sensor.

the number of frames transmitted in the entire pilot sequence. The relationship of f_s, f_d, and M_p is shown as follows.

$$M_p = 2 \times \frac{2f_d}{f_s}. \tag{3}$$

The two types of pilot sequence images are shown in Fig. 6. From Fig. 6, we can see that when the LED lights are ON in the first half, the corresponding LEDs turn OFF in the second half. We use these two kinds of pilot images for the decoding process which is described in detail in Section IV.

IV. Principle of VLC Signal Reception Using Rolling Shutter

In this section, we first describe the acquisition mechanism of the received image when capturing a fast-blinking LED array using a rolling shutter image sensor in a moving environment. As mentioned above, demodulation requires that the received image be divided into lines. This section describes the details of the demodulation process, including the dividing process.

A. Features of Rolling Shutter Image Sensors

First, we described that signals can be received by the RS image sensor even in a mobile environment. A schematic diagram of the timing chart for capturing a high-speed blinking LED array is shown in Fig. 7. Let T_c, T_e, and T_r denote the shooting time of one frame, the exposure time per scan line, and the readout time, respectively. The number of scan lines exposed during the display of one transmission frame is denoted by L_n. A certain scan line of Fig. 7 is shown in Fig. 8. Here, T_s is

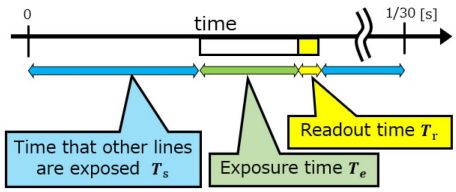


Fig. 8. Timing chart focusing on a single line in Fig. 7.

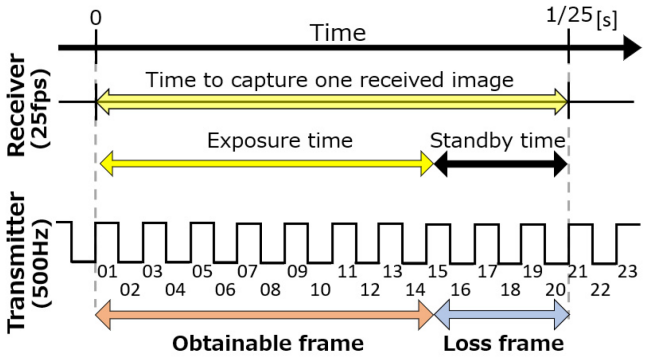


Fig. 9. Timing chart for capturing a single image with the RS image sensor at a shooting speed of 25 fps and an LED display speed of 500 Hz.

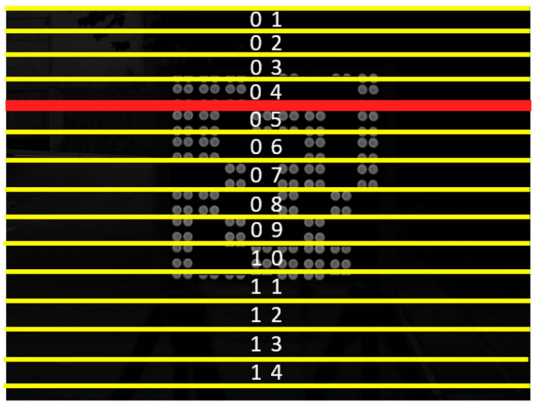
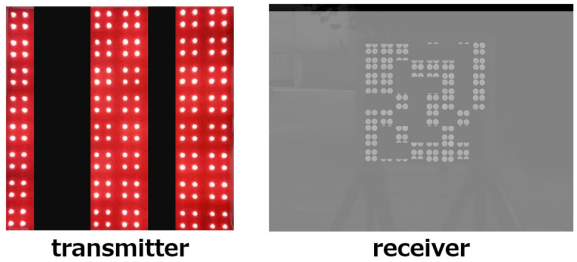
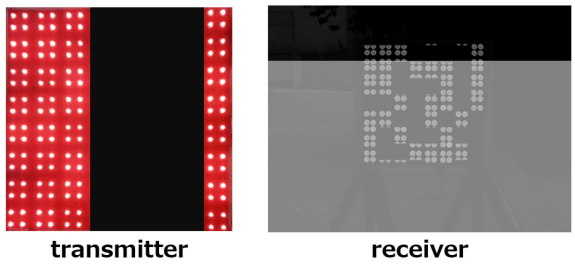


Fig. 10. Example of a received image of a data sequence.



(a) $m = 1$



(b) $m = 4$

Fig. 11. The m-th transmission frame and the range of received images to be captured by the time it is displayed.

the exposure time for lines other than the one shown here. The region of exposure time T_e shifts to the right as moving to the bottom line. The following relationship is established.

$$T_c = T_e + T_s + T_r. \tag{4}$$

The time T_s is subtracted from the exposure time T_c to obtain the actual exposure time of one line, which is as short as a 1/1,000-second scale. Therefore, the signal from the high-speed blinking LED array can be received by the RS image sensor with a low shooting speed.

B. Reception of LED Array Signal

We show the timing chart for capturing one image in Fig. 9 and an example of a received image of the data sequence in Fig. 10. Here, $f_s = 25$ fps, $f_d = 500$ Hz. The area between the adjacent yellow lines in Fig. 10 is the blinking pattern of the LED array displaying a transmission frame ($1/f_d$ [s]). For example, when the first transmission frame marked "01" in Fig. 9 is displayed (i.e., $1/f_d$ [s] after the start of shooting), the area up to the line marked "01" in Fig. 10 is captured. Next, when the second transmission frame marked "02" in Fig. 9 is displayed (i.e., $2/f_d$ [s] after the start of shooting), the area up to the line marked "02" in Fig. 10 is captured. This is because the RS image sensor exposes and takes a picture from the top to the bottom of the image sensor. For the sake of explanation, the timing of the display of the first transmission frame and the timing of the start of shooting are aligned here. Noteworthy, this timing could be shifted because the transmitter and receiver communicate asynchronously in practical applications.

Now, we specifically explain each scan line. First, we focus on the time period where the first transmission frame marked "01" is displayed on the transmitter side in Fig. 9. The transmission frame displayed at this time is shown on the left side of Fig. 11(a). Here, the data block $\boldsymbol{d}(i) = [1, 0, 0, 1, 1, 0, 1, 1]$ is transmitted. On the other hand, the receiver camera takes a picture of the line marked "01" in Fig. 10, as shown on the right side of Fig. 11(a). As can be seen from Fig. 11(a), the LED array is not captured in the line taken at this time. This means that the data block $\boldsymbol{d}(i) = [1, 0, 0, 1, 1, 0, 1, 1]$ is not acquired by the receiver.

Next, we focus on the time period when the 4th transmission frame marked "04" shown in Fig. 9 is displayed. The 4th transmission frame displayed is shown on the left side of Fig. 11(b). In the 4th transmission frame, the data block $\boldsymbol{d}(i) = [1, 1, 1, 0, 0, 0, 0, 1]$ is transmitted. The receiver takes a picture of the line marked "04" in Fig. 10, as shown on the left side of Fig. 11(b). Now, the receiver starts to capture the LED array part in the image. The signal is received in the line marked "04," meaning the data block $\boldsymbol{d}(i) = [1, 1, 1, 0, 0, 0, 0, 1]$ is acquired. In the above description, the 4th frame was captured

in line "04." However, the 4th frame is not necessarily captured in line "04" because the transmitter and receiver communicate asynchronously. If the start shooting time shifts slightly, the frame will be shot in another line such as 05 or 06. This is the reason why we vertically divide the transmitted frame. Although the captured line could fluctuate up and down, the received signal remains the same by dividing the transmission frame vertically.

Then, we focus on the moment when the LED array switches from the 4th to the 5th transmission frame. At this moment, the receiver side captures the red line between lines 04 and 05, as shown in Fig. 10. On the upper side of the red line, the data block $\boldsymbol{d}(i) = [1, 1, 1, 0, 0, 0, 0, 1]$ is captured, and on the lower side, the data block $\boldsymbol{d}(i) = [1, 1, 0, 1, 1, 1, 0, 1]$ is captured. To recover each data block, the receiver requires to divide the captured LED array according to the red line. In addition to the red line, there are several other areas that need to be divided, which are indicated by yellow lines. These dividing lines are called "boundary" in this study. For a certain LED display speed f_d and a certain camera's shooting speed f_s, the interval between adjacent boundaries is constant in a single captured image. Once the interval of this boundary is measured, it can be used to divide the captured LED array into the data block ranges. In this study, we obtain the length of the interval in advance and use it as known information. In addition, because the boundary interval is constant, we can determine the position of all boundaries when a reference boundary is identified. The following section describes the method used to identify the reference boundary in detail.

Next, we focus on the time period when the 11th transmission frame is displayed. At this time, the receiver is capturing the 11th line, as shown in Fig. 10. At this time, the entire LED array has been captured. In the captured image, the data blocks transmitted from the 4th to 11th transmission frames are received.

Now, we focus on the time period when the 15th transmission frame is displayed. At this time, all the lines of the first image are captured. However, it does not mean that the camera starts to capture the second image immediately. As shown in Fig. 9, there is a waiting period to read out the image after the exposure time. During this period, the camera does not expose. This means the receiver loses the signal transmitted from the 16th to the 20th transmission frames. The length of the waiting time varies depending on the kind of image sensor.

Finally, we focus on the time period when the 21st transmission frame is displayed. As shown in Fig. 9, the timing for displaying the 21st transmitted frame is the same as the timing for capturing the second received image. In other words, when the first transmission frame is displayed, the first line ("01") of the first received image is captured. When the 21st transmission frame is displayed, the same area as the first line ("01") of the second received image is captured, and so on. In short, the position of the boundary of the received image does not change until we stop shooting pictures.

C. Trimming the LED Array Points

In this section, we describe the method to trim out the portion of the LED array from the captured image. Because the intended communication environment is a public road, the ambient light reflected from guardrails and cars is captured as noise. In addition, the position of the LED array between successive frames fluctuates significantly in a moving environment. To demodulate the transmitting signal correctly, it is necessary to eliminate the noise and track the LED array from a single captured image.

Fig. 12. Example of captured image($P(j)$). Reflected light from cars, etc. is visible.

First, we set a short exposure time (T_e) to reduce the impact of the ambient light. An example of the captured image is shown in Fig. 12. As one can see, we eliminate most of the effects of ambient light. To remove the remaining noise, we use the first-order differential filter for the captured images. The process of first-order differential filtering is as follows. We first convert the captured image to a grayscale image. The pixel values of the converted image are represented by a two-dimensional matrix. Let $P(j)$ denote the matrix of the j-th captured image from the received image sequence. Let U and V denote the number of pixels in the vertical and horizontal directions, respectively. The grayscale value of the pixel in the u-th column and the v-th row is represented by $P_{u,v}(j)$, and the $P(j)$ matrix is represented as follows.

$$P(j) = \begin{pmatrix} P_{0,0}(j) & P_{1,0}(j) & \cdots & P_{U-1,0}(j) \\ P_{0,1}(j) & P_{1,1}(j) & \cdots & P_{U-1,1}(j) \\ \vdots & \vdots & \ddots & \vdots \\ P_{0,V-1}(j) & P_{1,V-1}(j) & \cdots & P_{U-1,V-1}(j) \end{pmatrix}. \tag{5}$$

Let $P_x(j)$ and $P_y(j)$ denote the horizontally differentiated and vertically differentiated matrices of $P(j)$, respectively. The u-th element from the top and the v-th element from the left of these two matrices are obtained as follows.

$$P_{x,u,v}(j) = P_{u+1,v}(j) - P_{u-1,v}(j) \tag{6}$$

$$P_{y,u,v}(j) = P_{u,v+1}(j) - P_{u,v-1}(j). \tag{7}$$

Let $P'(j)$ denote the edges matrix of the captured image $P(j)$, and $P'(j)$ is calculated as follows.

$$P'_{u,v}(j) = \sqrt{(P_{x,u,v}(j))^2 + (P_{y,u,v}(j))^2}. \tag{8}$$

The calculated edges image of the captured image is shown in Fig. 13. As shown in Fig. 13, after the first-order differential filtering, we almost remove the noise from the captured image. Finally, we search for the element in $P'(j)$ that is greater than a

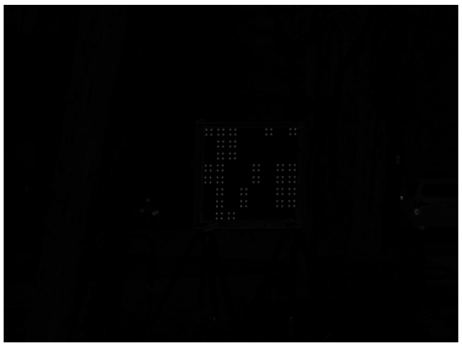


Fig. 13. Example of extracted edges from a captured image($P'(j)$). We can confirm that the noise has been eliminated.

threshold and record the coordinates. Let x_{left} and x_{right} denote the coordinates of the left and right edges, respectively. Let y_{top} and y_{bottom} denote the coordinates of the left and right edges, respectively. The image matrix with only the LED array part is denoted by $E(j)$, and $E(j)$ is expressed as follows.

$$E(j) = \begin{pmatrix} P_{x_{\text{left}},y_{\text{top}}}(j) & \cdots & P_{x_{\text{right}},y_{\text{top}}}(j) \\ \vdots & \ddots & \vdots \\ P_{x_{\text{left}},y_{\text{bottom}}}(j) & \cdots & P_{x_{\text{right}},y_{\text{bottom}}}(j) \end{pmatrix}. \quad (9)$$

D. Pilot Sequence Reception Image Identification

After tracking the LED array from the captured image, we use the pilot sequence to identify the location of the boundaries and detect the beginning position of the data sequence. To identify the images of the pilot sequence from all captured images, we generate a composite image $S'(j)$ by adding the pixel values of two captured images together and dividing by 2 as follows.

$$S'(j) = \frac{P(j) + P(j+2)}{2}. \quad (10)$$

Then, we generate $S(j)$ by trimming the LED array part from $S'(j)$ using the same process of the generation of $E(j)$. Finally, the smoothing process is performed to smooth the generated $S(j)$. The kernel K_b used in this process is defined as follows.

$$K_b = \frac{1}{a^2}\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \quad (a \times a \text{ matrix}), \quad (11)$$

where a indicates the size of the kernel, and a is calculated as follows.

$$a = \frac{x_{\text{right}} - x_{\text{left}}}{8}. \quad (12)$$

We show 6 consecutive received images in Fig. 14. First, we describe the cases $j = 1$ and $j = 2$. If one of the images, which is used to composite, is captured from the data sequence, the brightness of the smoothed image $S(j)$ will be uneven. An example of this result is shown in Fig. 15(a). However, when we composite two images captured from the pilot sequence, the luminance of the smoothed image $S(j)$ is flat. An example of this result is shown in Fig. 15(b). Based on this feature, the standard deviation of the pixel values of each image is calculated to identify the pilot sequence images.

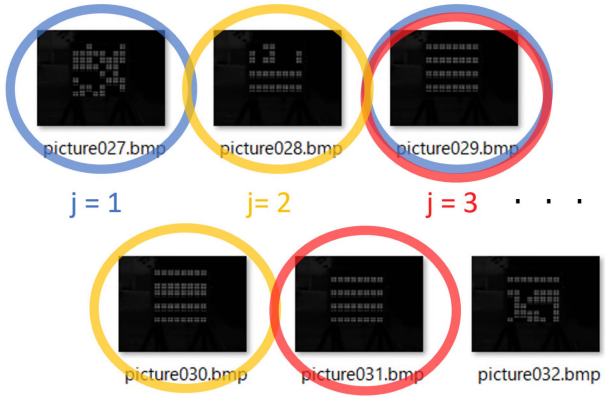
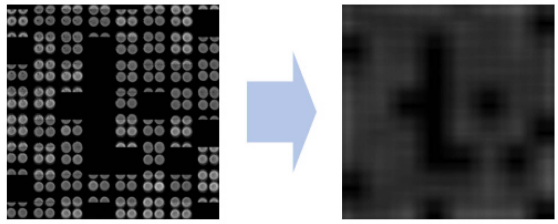
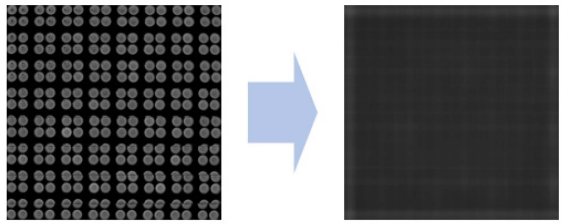


Fig. 14. How to combine images. The images are generated by matching the luminance by skipping one by one as shown in the circle.



(a) data sequence



(b) pilot sequence

Fig. 15. Left image is the composite image shown in Fig. 14, and the right image is the blurred image. The composite image of the pilot sequence shows no unevenness in luminance.

E. Segmentation of the LED Array in the Received Image

As we described above, the trimmed LED array is further divided into the ranges that represent 1 b of data. An example of dividing a data sequence image is shown in Fig. 16. In this study, we vertically divide the LED array into 8 columns to transmit 8 bits in a transmission frame. Therefore, the LED array captured in the image also needs to be divided into 8 equal parts. As described in Section IV, the boundary (horizontal yellow line) fluctuates up and down depending on the timing of the camera's shooting started. To locate the position of each boundary, a reference boundary is required to be identified from the captured image.

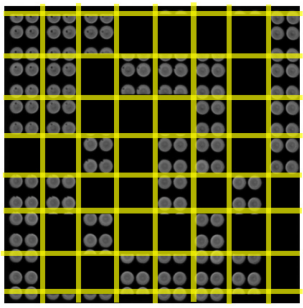


Fig. 16. Example of image segmentation for a certain data sequence.

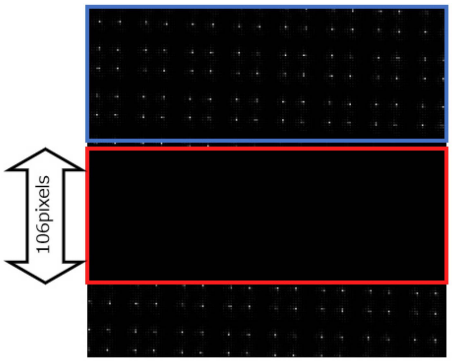


Fig. 17. Boundary identification methods. The luminance within the range is checked in order from the top to search for areas where dark lines are captured. Since there are no bright pixels in the red box, this is where the dark lines were taken, and the upper edge of the red box is used as the reference boundary.

The proposed method uses the captured images of the pilot sequence to identify the reference boundary. In the pilot sequence image, bright and dark lines are alternately captured at certain intervals. As an example, a portion of the pilot sequence image, when the LED display speed is set to 1,000 Hz, is shown in Fig. 17. In this image, the bright and dark lines are switched every 106 pixels, which is measured in the advanced process. The first step is to find where the dark rows were captured. We search for rows in the range of 1 to 106 pixels, and if a pixel is larger than a certain luminance value, then search for rows in the range of 2 to 107 pixels. As shown in Fig. 17, because bright pixels exist in the blue box, the search area is moving down. When the red box is searched, there are no bright pixels within the search area. Here, the upper edge of the red box is specified as the coordinates of the reference boundary.

F. Data Demodulation

The LED array in the received image is divided into a range of 1 b, as shown in Fig. 16. The average brightness is then calculated for each range, and the value between the highest and lowest values is used as the threshold value. The threshold value is calculated for each received image. Finally, we recover data for each range based on threshold discrimination and calculate the number of error bits. In this study, the lost transmission frames that are not captured in the image are not accounted for in the bit-error rate (BER) calculation.

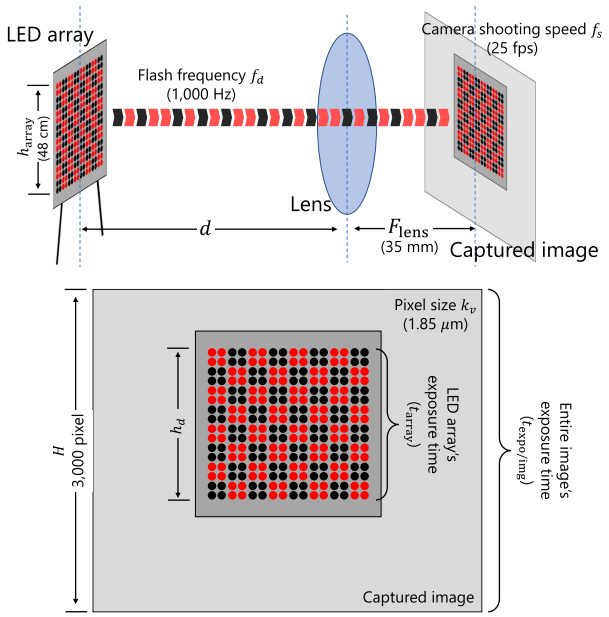


Fig. 18. Relationship between an LED array's captured area and its exposure time in the rolling-shutter-based VLC system.

V. Data Rate in Rolling-Shutter-Based Vehicle VLC

In this section, we perform a systematic analysis of the data rates of the proposed scheme. In this study, we divide the data rate into two categories: the data transmission rate and the data reception rate.

Here, we explain the data transmission rate first. In this study, a transmission frame (i.e., one blinking pattern of the LED array) transmits n bits of data and the LED array transmitter blinks transmission frames at f_d [Hz]. Let R_{trans} [bps] denote the data transmission rate, R_{trans} is calculated as follows:

$$R_{\text{trans}} = n \cdot f_d. \tag{13}$$

The R_{trans} depends on the flash frequency of the LED array transmitter and the number of bits transmitted in a transmission frame. The value of R_{trans} is not related to other factors.

In the rolling-shutter-based VLCs, the data reception rate varies at different communication distances. This is mainly because the area of the LED array captured in the image varies at different distances. Considering an image is captured at a certain time (i.e., the camera's shooting speed f_s is fixed), the different captured LED array areas mean the exposure time of the LED array in the image is different. As shown in Fig. 18, assume that the vertical length of the LED array captured at d [m] is h_d [pixel]. The height of the entire captured image is H [pixel]. Let t_{array} denote the captured LED array's exposure time. Let $t_{\text{expo/img}}$ denote the exposure time of the entire image. The relationship between t_{array} and $t_{\text{expo/img}}$ are expressed as follows:

$$\frac{h_d}{H} = \frac{t_{\text{array}}}{t_{\text{expo/img}}}. \tag{14}$$

The LED array switches transmission frames at f_d [Hz], meaning n bits of data takes f_d^{-1} [s] to send. During the LED array's

exposure time (t_{array}), the number of transmission frames that the camera could capture is $t_{\text{array}}/f_d^{-1}$. For objective purposes, we round down the value of $t_{\text{array}}/f_d^{-1}$, and calculate the received number of bits in the image, which is denoted by $N_{\text{bits/img}}$, as follows:

$$N_{\text{bits/img}} = \left\lfloor \frac{t_{\text{array}}}{f_d^{-1}} \right\rfloor \cdot n = \left\lfloor \frac{h_d \cdot t_{\text{expo/img}}}{H} \cdot f_d \right\rfloor \cdot n. \quad (15)$$

For example, if $t_{\text{array}} = 10.1$ [ms] and $f_d^{-1} = 1$ [ms], we approximate the number of received transmission frames in the captured image to be 10 ($=\lfloor 10.1 / 1 \rfloor$).

The camera's shooting time (T_c) of a frame equals the entire exposure time ($t_{\text{expo/img}}$) plus the frame's readout time (T_r) as follows:

$$T_c = t_{\text{expo/img}} + T_r = f_s^{-1}. \quad (16)$$

Let R_{rec} denote the data reception rate at d [m], and R_{rec} is calculated as follows:

$$R_{\text{rec}} = N_{\text{bits/img}}/T_c = \left\lfloor \frac{h_d \cdot t_{\text{expo/img}}}{H} \cdot f_d \right\rfloor \cdot n \cdot f_s. \quad (17)$$

In the experiment, f_d, $t_{\text{expo/img}}$, H, and f_s are constant parameters, therefore the data reception rate (R_{rec}) varies depending on the parameter h_d. Let h_{array} [m] denote the physical vertical length of the LED array. Let F_{lens} [mm] denote the focal length of the camera's lens. We calculate h_d [pixel] as follows:

$$h_d = \frac{h_{\text{array}} \cdot F_{\text{lens}}}{d \cdot k_v}, \quad (18)$$

where k_v [μm] denotes the vertical size of a pixel. Taking (18) into (17), we have the relationship between the data reception rate (R_{rec}) and the communication distance (d) as follows:

$$R_{\text{rec}} = \left\lfloor \frac{h_{\text{array}} \cdot F_{\text{lens}} \cdot t_{\text{expo/img}}}{d \cdot k_v \cdot H} \cdot f_d \right\rfloor \cdot n \cdot f_s. \quad (19)$$

Finally, the relationship between the data reception rate (R_{rec}) and the communication distance (d) is summarized in (19).

VI. Vehicular VLC Experiment

In this study, we conducted the roadside-to-vehicle communication experiment to evaluate the performance of the proposed system model. The specifications and results of the experiment are described in this section.

A. Experimental Specifications

Here, we describe the experimental setup. The experimental specifications are listed in Table I. We conducted the experiments in outdoor environments as shown in Fig. 19. In this experiment, the LED array transmitter was set in front of the forward direction of the vehicle. We faced the LED array's light-emitting part to the receiver camera's lens in the moving vehicle. The vehicle was driven straight toward the LED array transmitter and received VLC signals using an RS image sensor. The LED array has 16 × 16 LEDs arranged in a matrix and the size of the light-emitting matrix is 48 cm × 48 cm. We vertically divided the LED array into 8 columns and transmit 8 bits of data using a transmission frame (i.e., 8 bits in a blinking pattern). The blinking frequency of the LED array was set to 1,000 Hz.

TABLE I
Experimental Specifications

Experimental environment	Outdoor
Transmitter	a 16 × 16 LED array
Size of the LED array	48 cm × 48 cm
Number of bits a frame transmits (n)	8 bits
Blinking frequency (f_d)	1,000 Hz
Receiver Camera	U3-3890SE-C-HQ
Image sensor resolution	3,000 × 4,000
Shutter type	Rolling shutter
Pixel size	1.85 μm
Camera shooting speed (T_c)	25 fps
Exposure time	40 μs
Focal length	35 mm
Vehicle speed	10, 15, 20, 25, and 40 km/h
Communication distance	5.0–70.0 m

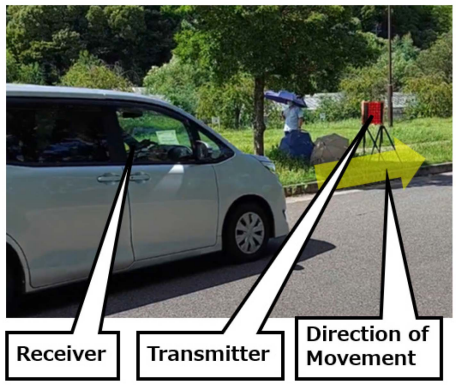


Fig. 19. Experimental scenery.

The receiver camera used in this study is IDS peak U3-3890SE-C-HQ with a 3,000 × 4,000 RS-type image sensor. The pixel size of the image sensor is 1.85 μm. We set the camera's shooting speed (f_s) to 25 fps and the exposure time per scan line (T_e) to 40 μs. The focal length (F_{lens}) of the camera lens that we used was 35 mm. During the experiment, the receiver camera is manually held in the vehicle. We drove the vehicle at a constant speed toward the LED array and captured the blinking of the LED array. Before the general roads experiment, we first conducted test experiments on a closed road section at the Nagoya University Higashiyama Campus and successfully received the VLC signals. Due to the speed limitation at the campus, the vehicle speeds (a.k.a receiver speeds) in test experiments were set to 10, 15, 20, and 25 km/h. To achieve the research target of this study, we then conducted the VLC experiment on the urban ordinary road. The speed limitation of the experiment road is up to 40 km/h. Because the feasibility of the proposed scheme at low vehicle speeds was confirmed and to circumvent the safety issues associated with driving at turtle

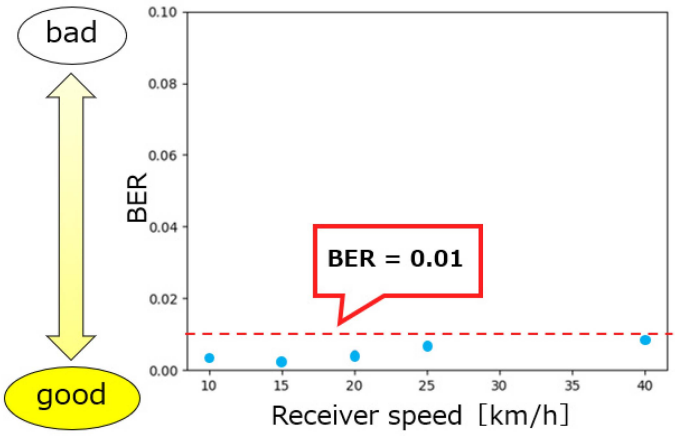


Fig. 20. Relationship between receiver speed and BER.

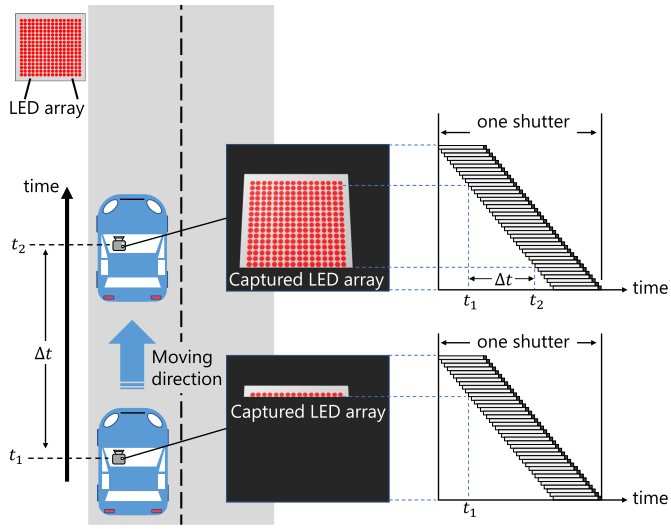


Fig. 21. Capture the LED array during the experiment.

speeds on general roads, we conducted the VLC experiment at a constant speed of 40 km/h. The experiment was conducted during the daytime and the weather was sunny. To consider the impact of the vehicle vibration and surrounding environment changes, all the experiments were conducted multiple times at the same vehicle speed. Finally, we measured the average BER at each speed to evaluate the demodulation performance of the proposed system.

B. Communication Performance Evaluation

We illustrate the BER plotted against the receiver speed in Fig. 20, revealing that the measured BER ranges from 0.002 to 0.008, regardless of the vehicle's speed. In this study, the transmitted signals were solely modulated by OOK. We believe that demodulation performance could be improved by incorporating error-correcting code such as convolutional code [26]. In addition, the decoding accuracy could be improved by using the normalization operation for the received LED luminance, which could be referred to in [27].

Let us now delve into the reasons for the demodulation errors. We believe that there are two causes of bit errors. The first is the LED array's inadequate trimming or division. To analyze the demodulation errors that occur due to inadequate trimming or division of the LED array, we model the experimental scenario as shown in Fig. 21. In this study, we consider the case where the vehicle drives straight toward the LED array transmitter. As shown in Fig. 21, the camera captures the first row and the last row of the LED array with a time interval (Δt [s]) due to the RS effect. Let S_{vehicle} [m/s] denote the driving speed of the vehicle. The vehicle moves $S_{\text{vehicle}} \cdot \Delta t$ [m] forward during the exposure time of the LED array. In this case, the distance between the LED array and the camera gradually decreases, resulting in the area of the LED array captured on the image gradually increasing. As a result, the captured LED array becomes a trapezoid shape as shown in Fig. 21. With increasing the driving speed S_{vehicle}, the camera captures more severe distortion of the LED array. Because the receiver trims and divides the LED array based on the assumption that the captured LED array was square, the distortion of the captured LED array could cause the division error, thus negatively affecting the demodulation result. A potential solution to address the inadequate trimming or division problem is to incorporate perspective transformation into the decoding process. The perspective transformation (a.k.a Homography) is related to the change of the object in different shooting perspectives. It can transform the distorted trapezoidal LED array into its original square shape, thus improving trimming accuracy and increasing demodulation performance.

The second cause of demodulation errors is the suboptimal threshold value optimization. This is also mainly because of the inadequate trimming or division problem. In this study, we calculate the average brightness for each area that represents 1 b of data as shown in Fig. 16. The trimming or division issue makes the detected 1-bit area and the calculated average brightness value inaccurate, thus causing the demodulation error. The potential solution to this issue is to improve the trimming and dividing accuracy by incorporating perspective transformation as described above.

The proposed method is expected to achieve V2X VLC when the vehicle speed is further increased. As mentioned above, the RS image sensor takes a long time to capture an image. When the receiver speed exceeds 40 km/h, the position of the LED array changes significantly in consecutively captured images. To address this problem, the proposed method identifies and trims the position of the LED array from a single captured image. We also set the exposure time per scan line to a short value (40 μs) as a solution to receive VLC signals in high-speed moving environments. Thanks to the high signal-to-noise ratio characteristics of the VLC system, setting a short exposure time could significantly reduce the impact of the background in a high-speed mobile setting. We consider the optimization for the exposure time in the proposed vehicle VLC systems as future work. In this way, our system is expected to remove noise and

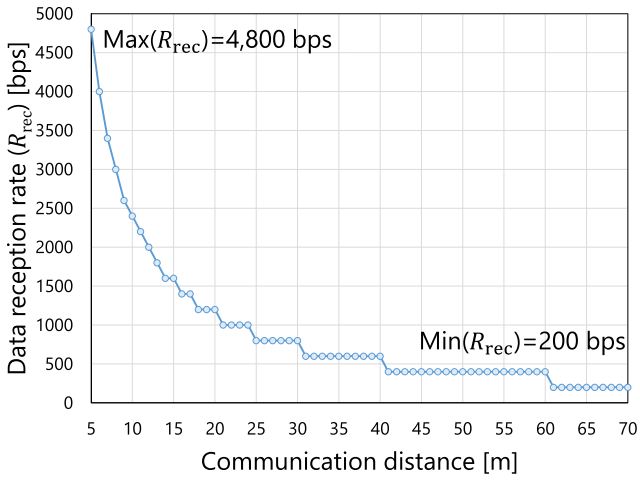


Fig. 22. Communication distance plotted against data reception rate in our experiment.

track the LED array transmitter correctly when the vehicle speed is 50 or 60 km/h.

Finally, we provide some information about the data reception rate (i.e., the maximum/minimum data rate) in the experiment. We measured the data reception rate ($R_{\rm rec}$) at different communication distances (d) using the experimental parameters listed in Table I. Assume that the readout time (T_r) shown in (16) is small enough to be neglected, the exposure time for an entire image ($t_{\rm expo/img}$) is considered to be approximately equal to the camera's shooting time (T_c). In the experiment, the communication distance ranges from 5.0 m to 70.0 m. Taking these parameters into (19), we have the relationship between $R_{\rm rec}$ and d, as shown in Fig. 22. As one can see, the maximum data reception rate ($R_{\rm rec}$) is achieved when $d = 5.0$ m, and $R_{\rm rec}$ reaches its minimum value when d is over 61.0 m. The maximum and minimum values of $R_{\rm rec}$ in this experiment are 4,800 bps and 200 bps, respectively.

VII. Conclusion

This paper presents a V2X VLC technique that utilizes an RS image sensor to operate in dynamic settings. The experimental results demonstrate that we were able to establish a VLC link between a roadside LED array and a vehicle traveling at 40 km/h. Due to the variability in the vehicle speed and the shooting speed, the positional relationship between the transmitter and receiver fluctuates significantly in the images captured continuously. As a solution, the proposed system tracks the LED array component in each captured image. Additionally, the short exposure time per line of the RS image sensor is utilized to detect signals from high-speed blinking LEDs in mobile settings. The proposed approach was verified in roadside-to-vehicle communication experiments with a BER on the order of 10^{-3}. We anticipate that this method can be applied to most urban roads. Future research will concentrate on enhancing the tracking precision of the LED array transmitter in high-speed mobile settings.

Acknowledgment

The authors would like to thank Prof. Masaaki Katayama (Nagoya Univ.), Prof. Hiraku Okada (Nagoya Univ.), Prof. BEN NAILA Chedlia (Nagoya Univ.) for useful discussions and suggestions.

References

[1] L. Zhu, F. Yu, Y. Wang, B. Ning, and T. Tang, "Big Data analytics in intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 383–398, Jan. 2019.

[2] N. Kumar, N. Lourenco, D. Terra, L. N. Alves, and R. L. Aguiar, "Visible light communications in intelligent transportation systems," in *Proc. 2012 IEEE Intell. Veh. Symp.*, 2012, pp. 748–753.

[3] 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 *Proc. IEEE GLOBECOM Workshops*, 2012, pp. 1234–1238.

[4] D. Iwase, M. Kasai, T. Yendo, S. Arai, T. Yamazato, and H. Okada, "Improving communication rate of visible light communication system using high-speed camera," in *Proc. IEEE Asia Pacific Conf. Circuits Syst.*, 2014, pp. 336–339.

[5] T. Yamazato et al., "Vehicle motion and pixel illumination modeling for image sensor based visible light communication," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1793–1805, Sep. 2015.

[6] Y. Goto et al., "A new automotive VLC system using optical communication image sensor," *IEEE Photon. J.*, vol. 8, no. 3, Jun. 2016, Art. no. 6802716.

[7] T. Yamamoto, T. Yamazato, H. Okada, M. Kinoshita, and K. Kamakura, "A comparison of distance performances of modulation schemes in an ITS image sensor communication," in *Proc. ICETC*, 2020, Paper B4-1.

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

[9] T. Hayashi, H. Okada, K. Kobayashi, T. Wada, and M. Katayama, "A study on visible light communications using a high-speed display and a rolling-shutter camera," in *Proc. Int. Symp. Nonlinear Theory Appl.*, pp. 316–319, Nov. 2020.

[10] R. Hamagami, T. Ebihara, N. Wakatsuki, and K. Mizutani, "Underwater visible light communication using phase-shift keying and rolling-shutter effect," in *Proc. IEEE 9th Glob. Conf. Consum. Electron.*, 2020, pp. 842–843.

[11] W. Wang, C. Chow, C. Chen, H. Hsieh, and Y. Chen, "Beacon jointed packet reconstruction scheme for mobile-phone based visible light communications using rolling shutter," *IEEE Photon. J.*, vol. 9, no. 6, Dec. 2017, Art. no. 7907606.

[12] T. Do and M. Yoo, "Visible light communication based vehicle positioning using a rolling shutter CMOS sensor," in *Proc. Int. Conf. Ubiquitous Future Netw.*, 2016, pp. 48–50.

[13] T. Zinda and W. Chujo, "Symbol synchronization method of image-sensor visible light communication using rolling shutter," in *Proc. IEEE 6th Glob. Conf. Consum. Electron.*, 2017, pp. 1–2.

[14] C. Liang, L. Chang, and H. Chen, "Analysis and compensation of rolling shutter effect," *IEEE Trans. Image Process.*, vol. 17, no. 8, pp. 1323–1330, Aug. 2008.

[15] Panasonic, "LinkRay," [Online]. Available: https://connect.panasonic.com/jp-ja/products-services_linkray

[16] T. Yamazato and H. Okada, "Practical applications of visible light communications," *J. IEICE*, vol. 101, no. 1, pp. 59–65, Jan. 2018.

[17] Y. Yang, F. Zhang, Z. Zeng, J. Cheng, and C. Guo, "Visible light communication for vehicular applications: A novel architecture with proof-of-concept prototype," *China Commun.*, vol. 20, no. 6, pp. 249–259, Jun. 2023.

[18] M. Y. Abualhoul, M. Marouf, O. Shagdar, and F. Nashashibi, "Platooning control using visible light communications: A feasibility study," in *Proc. IEEE 16th Int. Conf. Intell. Transp. Syst.*, 2013, pp. 1535–1540.

[19] A. Memedi and F. Dressler, "Vehicular visible light communications: A survey," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 161–181, Firstquarter 2021.

[20] T. Yamazato et al., "Image-sensor-based visible light communication for automotive applications," *IEEE Commun. Mag.*, vol. 52, no. 7, pp. 88–97, Jul. 2014.

[21] T. Do and M. Yoo, "Rolling shutter compensation for vehicle to vehicle positioning using CMOS sensor camera," in *Proc. Int. Conf. Inf. Commun. Technol. Convergence*, 2018, pp. 919–921.

[22] R. Hamagami, T. Ebihara, N. Wakatsuki, Y. Maeda, and k. Mizutani, "Rolling-shutter sensor-based visible light communication with cross-screen filter: Communication and positioning system using a commercial camera," in *Proc. IEEE 10th Glob. Conf. Consum. Electron.*, 2021, pp. 386–390.

[23] T. Nguyen, C. Hong, N. Le, and Y. Jang, "High-speed asynchronous optical camera communication using LED and rolling shutter camera," in *Proc. Int. Conf. Ubiquitous Future Netw.*, 2015, pp. 214–219.

[24] T. Zinda, K. Ito, and W. Chujo, "Rolling-shutter-based optical camera communication using distributed LED array," in *Proc. Int. Symp. Commun. Syst., Netw. Digit. Signal Process.*, 2018, pp. 1–4.

[25] T. Arisue et al., "BER measurement for transmission pattern design of ITS image sensor communication using DMD projector," in *Proc. IEEE 17th Annu. Consum. Commun. Netw. Conf.*, 2020, pp. 1–6.

[26] A. Neubauer, J. Freudenberger, and V. Kuhn, "Convolutional Codes," in *Coding Theory: Algorithms, Architectures and Applications*. Hoboken, NJ, USA: Wiley, 2007, pp. 97–161.

[27] Z. Tang, S. Arai, and T. Yamazato, "Simplified alamouti-typespace-time coding for image sensor communication using rotary LED transmitter," *IEEE Photon. J.*, vol. 14, no. 1, Feb. 2022, Art. no. 7307007.