IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 18, 2019, accepted June 27, 2019, date of publication July 11, 2019, date of current version July 31, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2928417

Square Wave Quadrature Amplitude Modulation for Visible Light Communication Using Image Sensor

SYUKRON ABU ISHAQ ALFAROZI[1], (Member, IEEE), KITSUCHART PASUPA[1], (Senior Member, IEEE), HIROMICHI HASHIZUME[2], (Member, IEEE), KUNTPONG WORARATPANYA[1], (Member, IEEE), AND MASANORI SUGIMOTO[3], (Member, IEEE)

[1]Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
[2]National Institute of Informatics, Tokyo 101-8430, Japan
[3]Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan

Corresponding author: Kuntpong Woraratpanya (kuntpong@it.kmitl.ac.th)

This work was supported in part by the King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand, and in part by the ASEAN University Network/Southeast Asia Engineering Education Development Network (AUN/SEED-Net) under the Japan International Cooperation Agency (JICA), Japan.

ABSTRACT Most visible light communication (VLC) technologies use a light emitting diode (LED) as a data transmitter and a photodiode as a receiver. In this paper, we alternatively focus on the use of an image sensor or camera as a receiver due to its wide availability. However, the successful use of an image sensor mainly depends on the efficiency of the encoder-decoder and the modulation scheme. Thus, this paper proposes a novel modulation scheme based on a square wave signal called a square wave quadrature amplitude modulation (SW-QAM) method. This method can accommodate different camera settings and overcome the problem of LED flicker that is generally sensed by human eyes when the LED frequency is low. At the transmitter side, multiple LEDs can be used to increase the transmission bit rate, while, at the receiver side, a *Wiener* filter is used as a complementary technique to SW-QAM for solving the light interference phenomenon due to the closeness of one LED to another. Our experimental results show that the proposed SW-QAM scheme can decode symbols very well either the for close or far communication distances, dark or bright lighting conditions, and single or multiple LED points.

INDEX TERMS Visible light communication (VLC), image sensor communication (ISC), exposure time, square wave quadrature amplitude modulation (SW-QAM).

I. INTRODUCTION

Visible light communication (VLC) is a new research trend for wireless communication systems that use a light emitting diode (LED) as a transmitter (Tx) [1], [2]. VLC refers to optical wireless communication (OWC) systems of which the transmitter and receiver (Rx) are aligned along the line-of-sight (LOS) to each other. For the transmitter, the data are transmitted based on its LED luminance and modulation scheme. The use of LED as a communication component leads to a new breakthrough for OWC systems; that is, LED is used not only for lighting purpose but also for data communication. Moreover, the recent development of LED has made it extremely energy efficient and its production cost low. Regarding the receiver, there are two types of sensors that can be used as a VLC receiver: a photodiode (photodetector) and a matrix of photodiodes (image sensor or camera). Hereafter, the terms 'image sensor' or 'camera sensor' or 'camera' will be used interchangeably. The advantage of using a camera is that we can transform any mobile devices with a camera into a VLC receiver. This potentiality can enable anyone who has a modern mobile device to benefit from VLC technology.

Responding to this transformation, four important factors need to be considered in developing a modulation scheme

The associate editor coordinating the review of this manuscript and approving it for publication was Mamoun Alazab.

based on LED as transmitter and camera as receiver for VLC. These factors are (i) error rate of symbol/signal decoding, (ii) communication speed between transmitter and receiver (bit per second: *bps*), (iii) LED flicker problem of the transmitter, and (iv) flexibility of parameter settings of the modulation scheme. Especially in the last factor, this flexible parameter setting opens the system for using different LED frequencies and camera speeds (frame per second: *fps*). The following paragraph briefly describes the focus of the existing papers in VLC.

Most of the works in VLC development are based on photodiode receivers such as [3]–[5] due to its high communication speed. On the other hand, cameras or image sensors, bundled with mobile devices, can easily be transformed into communication receivers. However, there are some challenging limitations of an image sensor in VLC applications. For example, an image sensor usually provides a lower sampling rate due to a large number of photodiodes. This large number of photodiodes is also a cause of a lower frame rate, typically below 100 *fps*. This limitation leads to low data rate communication. Even though most of the recent works [6]–[8] attempted to improve the communication performance by using an on-off-keying-based (OOK-based) modulation scheme, they still mostly provided a low bit rate. One way to increase the bit rate of OOK is by using a high sampling frequency. Therefore, a rolling shutter camera can be a solution since it can provide a sampling frequency more than 100 times faster than its frame rate by using the line sampling technique, as reported in [9], [10]. However, a high sampling frequency and rolling mechanism may cause interference between pixel lines, especially when the luminance of the light is not evenly distributed. This undesirable phenomenon can be solved by using an adaptable thresholding method for discriminating between 1 (ON) and 0 (OFF) symbols. LED flicker problem is also one of the focus factors in developing a modulation scheme for VLC. A flicker problem occurs when the blinking frequency of the LED is less than 100 Hz [11]. The different solutions of the flicker problem were reported in [9], [12]–[15]. Here we can see that the communication speed and LED flicker problem are the main focus of existing research in VLC, but the flexibility of parameter settings of the modulation scheme is not a focus point of those research papers.

Inspired by Sugimoto's work [16] that was recently proposed a time synchronization technique for optimally modulated illumination, we extend his work by using the time difference between the camera and the modulated signal as a phase of the signal. Therefore, we can utilize both the phase and amplitude for signal modulation. With this ability, a novel scheme of quadrature amplitude modulation (QAM) called a square wave quadrature amplitude modulation (SW-QAM) for VLC technologies based on an image sensor is proposed for the achievement of a very low error rate, fast communication speed, no LED flicker problem, and easy parameter settings. Note that Sugimoto's work was proposed for the time synchronization technique, but not for the communication scheme.

The main contributions of our work in the proposed SW-QAM method are listed as follows.

- A novel modulation scheme, called SW-QAM, is formulated by a generic and simple mathematical expression. In practice, it can adopt different settings of exposure time ratio η and different frequencies of symbols f_s (see full details of η and f_s in Section III), thus making it flexible to a variety of communication environments.
- SW-QAM supports flicker-free modulation when f_s is set to a high frequency of at least 100 Hz.
- Preamble and frame-overlapping detection methods are described for determining the starting block of symbols. The accurate time difference can identify phase shifting or rotation of QAM constellation, thus the time difference between the rising edge of the camera and the starting edge of the symbol can be calculated correctly.
- Multiple LEDs scheme at the transmitter side is investigated to increase the communication speed.

The rest of this paper is organized as follows. Section II briefly reviews the existing studies of VLC technologies and some modulation techniques that are used in image sensor communication (ISC) and addresses problem statement as well. Section III explains the proposed SW-QAM scheme and other essential techniques including preamble detection and frame discarding. Then, the experimental results are presented and discussed in Section IV, and finally, we conclude our work in Section V.

II. BACKGROUND AND PROBLEM STATEMENT

Usage of image sensors is another growing trend in VLC research area even though the communication speed of VLC with image sensor is low when compared to that of VLC with a single photodetector, due to the sampling rate of VLC with image sensor is constrained by its frame rate (frame per second: *fps*). The VLC with image sensor scheme is more beneficial because of being not only for image capture but also for data communication, as reported in [17], [18]. In VLC research area, the modulation scheme and the transmitter/receiver are interdependent. In other words, it is difficult to develop a modulation scheme that fits all kinds of transmission/receiver devices. An efficient modulation scheme depends on recent advanced transmission/receiver devices that are key to the success of the system.

In this section, we briefly review the existing VLC-based communication methods that are based on two types of cameras: rolling shutter camera (high speed camera) and global shutter camera (low speed camera). The most convenient and simple scheme for VLC is on-off-keying (OOK). There are various implementations of OOK scheme. For instance, the works in [12], [13] used an A-On-A′-Off protocol with oversampling of n times repeated data. In their works, they used a high speed camera, rolling shutter camera, with 1,000 fps to mitigate flicker problem. A rolling shutter

camera has its own advantages such that its sampling rate can be increased through the rolling mechanism by sequentially exposing pixel lines on the captured image. Some existing OOK communication schemes [9], [14], [15] utilized this kind of rolling shutter camera to send a symbol by exposing every pixel line as 1 or 0 symbols. However, the Tx and Rx devices must be close enough so that the system can provide a large area of the captured light source. As the rolling shutter camera exposes the light from every pixel line, the lines not covered by the light become missing symbols. As a result, the amount of data captured by the rolling shutter camera is proportional to the size of the light source that covers the pixel lines. Another modulation scheme that takes advantage of the rolling mechanism is frequency shift keying (FSK), as reported in [19], [20]. This scheme satisfies the requirement of the *Nyquist* sampling theorem because the sampling rate of the rolling shutter camera is 2 times greater than the blinking frequency of LED. Theoretically, FSK is more reliable than amplitude modulation in terms of noise interference. However, it needs more pixel lines than the OOK scheme. In practice, FSK uses a single frame to represent a single frequency or symbol regardless of the frame overlapping symbol. This kind of FSK scheme also requires a large size of the light source, similar to the OOK scheme, in order to distinguish different symbols or frequencies. Recently, a modulation scheme based on a rolling shutter camera called a hybrid of M-FSK and 2-PSK (phase shift keying) [21] that provides a good performance was proposed. The system mainly utilized M-FSK modulation scheme that provides high data rate by using two line LEDs. However, when the camera cannot distinguish the frequencies at some distance, the system switches to use 2-PSK modulation scheme by detecting the phase difference between the two LEDs. By combining these two modulation schemes together, the hybrid of M-FSK and 2-PSK is more advantageous in terms of wider channel availability at different communication distances. On the other hand, development of a modulation scheme that supports low speed camera such as global shutter camera is also required in order to support a wide range of devices. The following paragraph briefly describes this kind of modulation scheme.

Other modulation schemes that support global shutter camera also have been developed. As for now, efficient global shutter cameras are being developed as a future technology [22]. For this reason, we believe that in the future global shutter camera will become standard in this field as a solution for the problem of rolling shutter distortion. An example of a recently developed scheme that supports the global shutter camera is a spatial 2-phase shift keying (S2-PSK) modulation scheme for vehicle-to-vehicle communication proposed by Nguyen et al. [6]. They utilized the neural network as a complementary method to localize LEDs in a captured image and to classify the captured image as bit 1 or bit 0 based on XOR decision boundary of two LEDs. In their work, the neural network was not mainly for encoding or decoding purposes but for LED localization method. However, they showed that by using the neural network as the decoding method, the system performed better. This S2-PSK scheme can also overcome flicker problem like other modulation schemes such as under-sampled frequency shift on-off-keying (UFSOOK) [8], under-sampled phase shift on-off-keying (UPSOOK) [7]. All of these modulation schemes use LED points instead of flooding the entire image with light. By using LED points in the image frame, communication over long distance under LOS constraint is possible. However, since the schemes use a global shutter camera, the bit rate that depends on a global shutter camera's frame rate is very slow.

To the best of our knowledge, a few papers [6], [21] introduced a total solution for VLC in a single scheme, but they did not report the flexibility of parameter settings and scalability of communication speed. Therefore, this paper proposes a novel method of SW-QAM scheme. This method is a new total solution that can achieve the VLC performance in terms of a very low error rate, fast communication speed, no LED flicker problem, and easy parameter settings. Furthermore, to evaluate the capability of the new total solution, the system is implemented with the following two challenges: (i) At the transmitter side, two LEDs are installed very close together whereas at the receiver side *Wiener* deconvolution—a simple image restoration technique—is used to remove light interference between two LEDs. (ii) Does the use of a low sampling rate camera such as global shutter camera still achieve fast communication speed or not? The successful proof of the former implies that the transmitter can be extended to apply for LED array, while the successful proof of the latter implies that the system can achieve a higher communication speed when implemented with a high sampling rate camera. The following Sections III and IV explain our proposed method and reveal the remarkable results, respectively.

III. PROPOSED METHOD

This paper proposes a novel modulation scheme, square wave quadrature amplitude modulation (SW-QAM), for VLC. The key concepts and implemented scheme are covered by the following topics: fundamentals of LED signal and image sensor; SW-QAM communication scheme; preamble signal, frame discarding, and preamble detection; possibility of various parameter settings of LED frequency f_s and camera *fps*, and *Wiener* deconvolution.

A. FUNDAMENTALS OF LED SIGNAL AND IMAGE SENSOR

In this subsection, the LED signal and image sensor are briefly described for easier understanding of our proposed scheme. The complete explanation of how a camera receives a signal from an LED by using a modulated square wave signal can be found in [16]. Here, as illustrated in Fig. 1, the camera typically has a built-in integration function along with its exposure time. The exposure time is the period of time that the camera's shutter is open to allow light in. During this time, the image sensor captures the cumulative luminance of the light coming in. Mathematically, without losing the

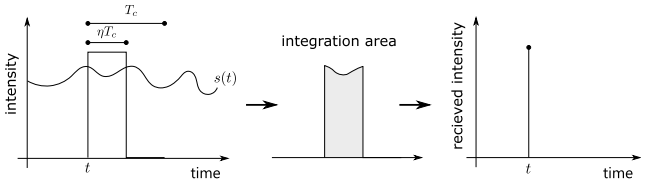


FIGURE 1. Integration sampling of an LED signal by an image sensor during its exposure time ηT_c at time t.

generality of non-square wave signals, if we have a signal from an LED in the form of intensity as a function of time $s(t)$ with a camera's captured period of T_c, and an exposure time of ηT_c, where $0 < \eta < 1$ is the exposure time ratio, then the light received by the camera at the shutter-release time t can be expressed by

$$r(t) = \int_{t}^{t+\eta T_c} s(t) \quad \mathrm{d}t. \tag{1}$$

According to (1), the camera integrates the signal during its exposure time of ηT_c. The result of this integration is recorded in the pixels of the image. We can extract the correct signal information from this recorded intensity in these pixels with an efficient encoding method and modulation scheme. In practice, the use of square wave signals for modulation is one of the most efficient methods as well explained in [16]. With a proper setting, we can obtain the time difference between the camera's shutter-release time and the rising edge of the LED signal. This time difference can be used as the phase of the LED signal. With this concept, we utilize not only the amplitude information but also the phase information from a modulated square wave signal.

B. SW-QAM COMMUNICATION SCHEME

The encoding and decoding stages can be described by a simple diagram as shown in Fig. 2. For easier explanation, important parameters are first introduced. Let A and T_s be the amplitude and period of a modulated square wave signal, respectively. In the encoding stage, a symbol is mapped onto a QAM constellation coordinate consisting of an amplitude A and phase θ. This symbol is modulated to a square wave signal with an amplitude A, a period T_s, and a shifting phase $t = T_s\theta/2\pi$, where t is in radian. Most importantly, the configuration of the encoder/decoder must satisfy **Definition 1**.

Definition 1: The configuration of T_s, T_c, and η is called a proper setting if the phase t and the amplitude A of the modulated signal is uniquely defined.

Here, we propose a proper setting, i.e., the modulated square wave period T_s must be sampled with three times of sampling frames by the camera with captured period T_c, $T_s = 3T_c$, and the exposure time ratio $\eta = 0.5$. With this setting, we can obtain both the amplitude and phase of a signal correctly. We note that the more sampling frames of the camera, the slower communication speed of the system; therefore, the setting of $T_s = 3T_c$ is the minimum number of sampling frames of T_s that satisfies **Definition 1** that will be explained in this subsection. In addition, other proper settings are proposed in the next subsection also based on 3-frame decoding scheme with different exposure time and symbol frequency. In the decoding stage, the camera sensor captures this modulated signal and decodes it by the SW-QAM decoding scheme to reconstruct the amplitude and phase of the original symbol. In this study, we provide a 3-frame decoding scheme for any proper settings. In other words, a modulated symbol is sampled in three consecutive frames by the camera sensor in order for the amplitude and phase of the original signal to be perfectly reconstructed. Other possible proper settings will be provided in the following subsection. One can flexibly use these proper settings based on the needs and purposes of the application.

From theory to practice, this paper proposes a simple mathematical expression for encoding (and decoding of) the amplitude and phase of transmitted symbols. For the encoder, we start with modeling a simple mathematical expression of frame intensity function. Let $r_1(t)$, $r_2(t)$, and $r_3(t)$ be the intensity functions of a modulated square wave with a phase shift t obtained from the first, second, and third frames of captured images, respectively. The intensity function of the first frame, $r_1(t)$, can easily be formed as depicted in Fig. 3. Mathematically, we have

$$r_1(t) = \begin{cases} r_{max} \cdot (1 - t/\eta T_c) & (0 \le t < \eta T_c) \\ 0 & (\eta T_c \le t < T_s/2) \\ r_{max} \cdot (t - T_s/2)/\eta T_c & (T_s/2 \le t < T_s/2 + \eta T_c) \\ r_{max} & (T_s/2 + \eta T_c \le t \le T_s), \end{cases} \tag{2}$$

where r_{max} is the maximum intensity among all three frames from the light source (LED-Tx) received by the camera sensor during the exposure time, which is also represented as the amplitude of the signal. Here, when $t = 0$ the integration area of signal is $A\eta T_c = r_{max}$, then it linearly decreases to 0 as $t = \eta T_c$. As t continuously increases from ηT_c to $T_s/2$, the integration area is constantly 0. Fig. 3 depicts the event as t increases further to T_s. In the same way, the intensity functions $r_2(t)$ and $r_3(t)$, which have different phases T_c, i.e., $T_s/3$ and $2T_c$, i.e., $2T_s/3$ from $r_1(t)$, respectively, are represented below

$$\begin{aligned} r_2(t) &= r_1(t - T_s/3) \\ r_3(t) &= r_1(t - 2T_s/3). \end{aligned} \tag{3}$$

Fig. 4(a) shows the modulation of the square wave signal based on amplitude A and phase t. Then, Fig. 4(b) shows a result of plotting the intensity functions $r_1(t)$, $r_2(t)$, and $r_3(t)$ when $\eta = 0.5$.

For the decoder, it is easy to model a mathematical expression for decoding the encoded symbol. With setting up $\eta = 0.5$ as illustrated in Fig. 4(b), the shifting phase t can be uniquely defined by the values of $r_1(t)$, $r_2(t)$, and $r_3(t)$ as expressed in (4).

$$t = \begin{cases} (i + 1 - r_{mid}/r_{max})T_s/6, & i \text{ is even} \\ (i + r_{mid}/r_{max})T_s/6, & i \text{ is odd}, \end{cases} \tag{4}$$

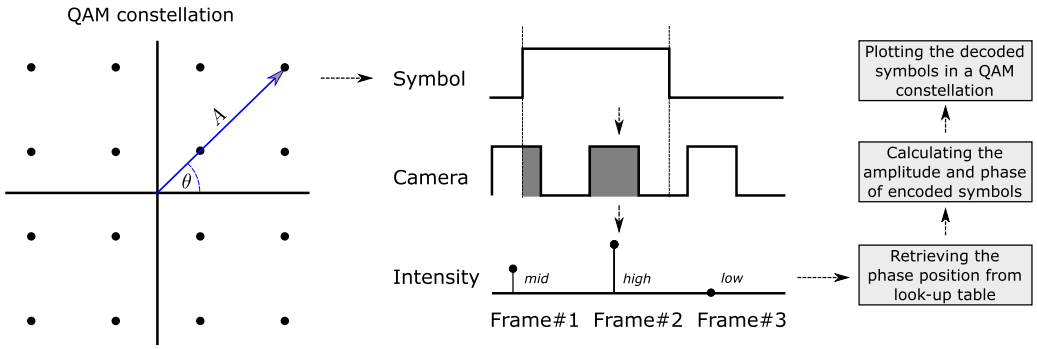


FIGURE 2. An overall framework of the proposed SW-QAM method: From QAM constellation encoding of a symbol to decoding amplitude and phase of the encoded symbol.

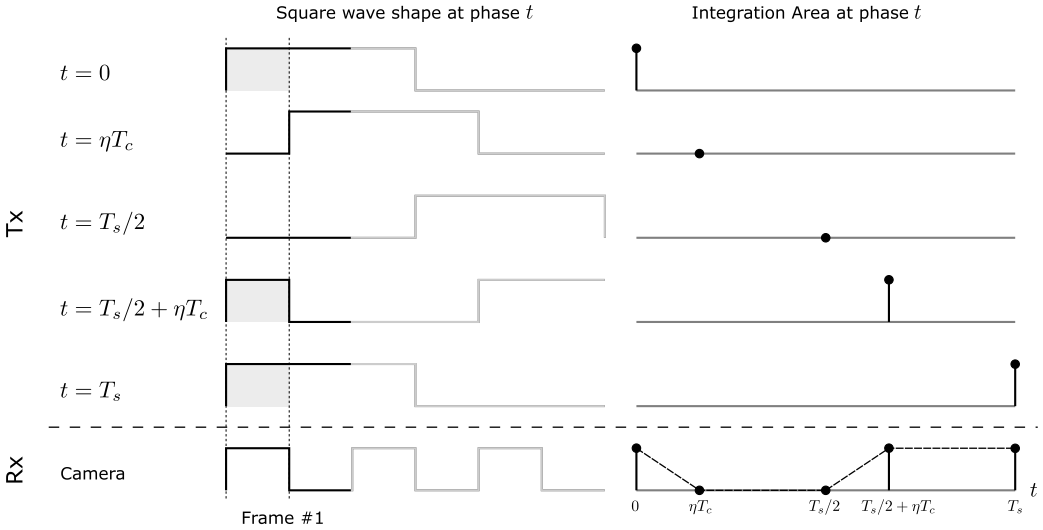
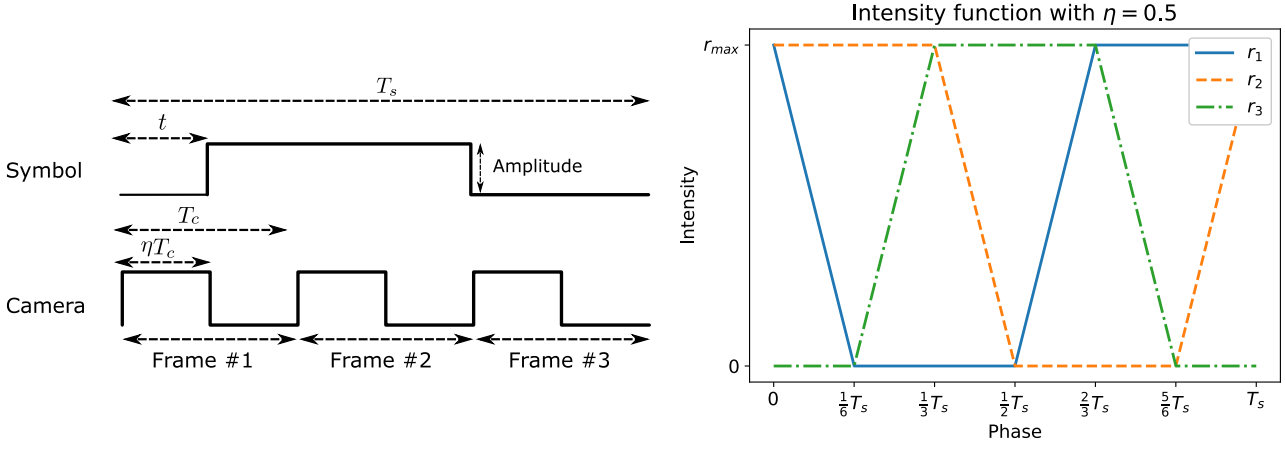


FIGURE 3. An illustration of how to determine the intensity function $r_1(t)$ of the Frame #1: At different phases, $t = 0$, ηT_c, $T_s/2$, $T_s/2 + \eta T_c$, and T_s of modulated square wave signal resulted five different integration areas, *max*, *low*, *low*, *max*, and *max* at the receiver, respectively, and the full line curve of a single period for the first frame intensity function shown at bottom right of this figure.



(a) (b)

FIGURE 4. An intensity function of each frame that depends on the phase t: (a) a symbol with a certain amplitude A and phase t is sampled by a camera with three consecutive frames, such that $T_s = 3T_c$, and (b) an overlapping plot of three intensity functions: $r_1(t)$, $r_2(t)$, and $r_3(t)$ of three frames: Frame #1, Frame #2, and Frame #3.

where i is the phase position interval of the intensity function obtained from Table 1, in which *low*, *mid*, and *high* are the lowest, medium, and highest intensity of frames, respectively. Let us refer to intensity functions $r_1(t)$, $r_2(t)$, and $r_3(t)$ in Fig. 4(b). For easy explanation, we only discuss the intensity function from $t = 0$ to $t = T_s$.

Here, each intensity function is divided into 6 intervals: $[0, 1/6T_s)$, $[1/6T_s, 1/3T_s)$, $[1/3T_s, 1/2T_s)$, $[1/2T_s, 2/3T_s)$, $[2/3T_s, 5/6T_s)$, and $[5/6T_s, T_s)$. The first to the sixth intervals are defined by the phase positions from 0 to 5, respectively. From this information, a simple decoding look-up table can be created by using the relation of the frame intensity function and its intensity value on each interval as shown in Table 1. In this table, the phase position i is mapped to the intensity value of such three intensity functions. Now, we start with the intensity function $r_1(t)$ at the phase position $i = 0$, which is the first interval $[0, 1/6T_s)$. The intensity value is linearly decreased from r_{max} to 0 as a slope line. This slope line is mapped as *mid*. Then, at the phase position $i = 1$, which is the second interval $[1/6T_s, 1/3T_s)$, the intensity value is constantly 0 as a low horizontal line. This low horizontal line is mapped as *low*. In the same way as done at the phase positions 0 and 1, the mapped intensities for the phase positions 2, 3, 4, and 5 are mapped as *low*, *mid*, *high*, and *high* according to a low horizontal line, a slope line, a high horizontal line, and a high horizontal line, respectively. For the intensity functions $r_2(t)$ and $r_3(t)$, we use the same way as $r_1(t)$ to map the intensity value of each interval to each phase position.

TABLE 1. Phase position interval *i* based on the intensity of the three frames.

i	$r_1(t)$	$r_2(t)$	$r_3(t)$
0	*mid*	*high*	*low*
1	*low*	*high*	*mid*
2	*low*	*mid*	*high*
3	*mid*	*low*	*high*
4	*high*	*low*	*mid*
5	*high*	*mid*	*low*

Theoretically, the *low* value of intensity must be zero for $(\eta T_c \leq t < T_s/2)$ according to (2). However, in real situations, the background intensity from the environment usually makes *low* not to be zero. Hence, to reduce this effect, the *low* value is subtracted from the *high* and *mid* values of the intensity as below,

$$\begin{aligned} r_{max} &= high - low \\ r_{mid} &= mid - low. \end{aligned} \tag{5}$$

Now, r_{max} and r_{mid} from (5) are ready for the phase shift calculation in (4) and for decoding the signal by using the SW-QAM method. Then, the reconstructed phase of symbols is converted to radian by the relation $\theta = 2\pi t/T_s$, so that it can be plotted in an SW-QAM constellation.

Furthermore, for $\eta \neq 0.5$ with $T_s = 3T_c$ as shown in Fig. 5, we cannot uniquely define the phase in any line intervals due to overlapped line segments. There are ambiguous areas in the reconstruction of the phase from the modulated signal. Hence, these configurations are called improper settings.

In practice, there are three problems that need to be addressed to implement SW-QAM: (i) non-synchronized setting between the LED signal and the camera sensor, (ii) frame overlapping, and (iii) LED flicker. The following subsection describes how to solve those problems and how to implement the proposed scheme in the real situation.

C. PREAMBLE SIGNAL, FRAME DISCARDING, AND PREAMBLE DETECTION

As mentioned in the previous Subsection III-B, the shifting phase in (4) is correctly defined if and only if the LED signal and camera sensor setting are synchronized. In our case, they do not need to be synchronized, because preamble signal is utilized to detect the time difference t_d between the rising edge of an LED signal and the shutter release of a camera. In our work, the preamble signal plays an important role in defining the starting block of symbols. As shown in Fig. 6, the packet consists of 8 symbols for preamble and 64 symbols for data. Here, the preamble symbol is a square

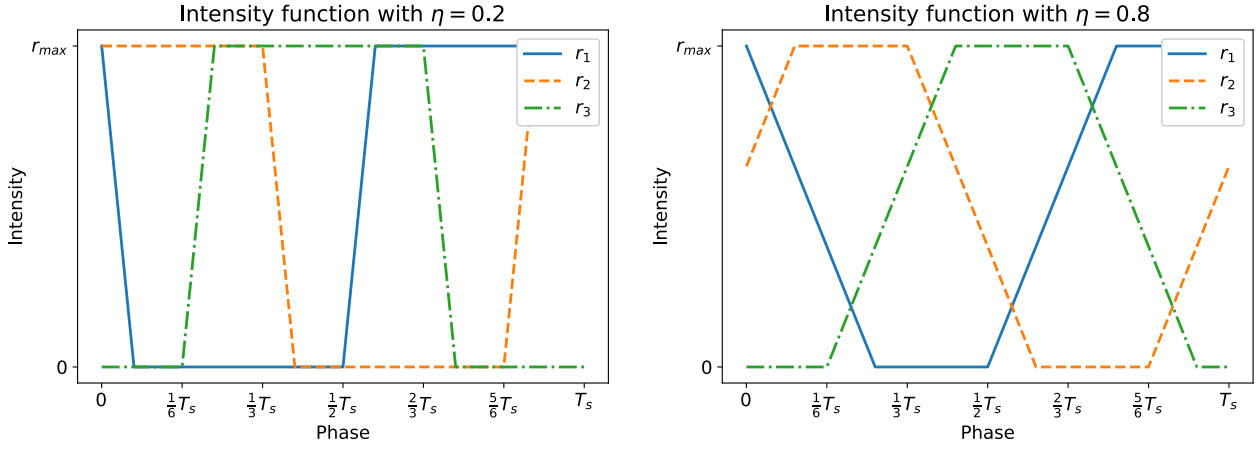


FIGURE 5. The intensity of each frame when the symbol has shifting phase *t* with base setting $T_s = 3T_c$ and different settings of η: (a) $\eta < 0.5$ and (b) $\eta > 0.5$.

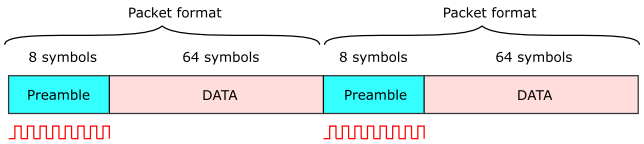


FIGURE 6. Packet format.

wave signal that has an amplitude of $A = 1$ and a phase of $t = T_s/2$ or $\theta = \pi$. Based on this format, the preamble symbol is sent 8 times before sending 64 data symbols. This preamble symbol can be utilized to calculate time difference t_d that is an additional phase shift to other symbols. Using this preamble information, the time difference t_d, due to non-synchronized setting, can be obtained by

$$t_d = T_s/2 - t_{pre}, \tag{6}$$

where t_{pre} is the average calculated phase of the preamble symbols obtained by camera. Thus, the shifting phase of other symbols, t_{sym}, is updated as follows,

$$t_{sym} := t_{sym} + t_d. \tag{7}$$

In addition, the amplitudes of these symbols are related to the amplitudes of the preamble signals. Hence, the normalized amplitudes of the symbols can be calculated as follows,

$$A_{sym} := A_{sym}/A_{pre}, \tag{8}$$

where A_{pre} is the average calculated amplitude of the preamble symbols. Fig. 7 graphically shows an non-synchronized setting on preamble symbols.

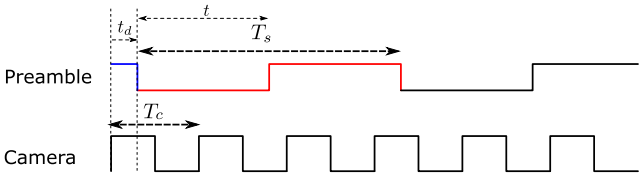


FIGURE 7. A non-synchronized setting to preamble symbols with a time difference between the camera and the symbol, t_d (different color means different part of symbol).

By **Definition 1** as described in Section III-B, three consecutive frames are sufficient to correctly decode a symbol. However, in practice, frame overlapping may occur under a non-synchronized setting. A frame overlapping is an event that a frame integrates different parts of symbols, thus leading to an incorrect integration. In order to solve this problem, a symbol has to be sent with a length of $4T_c$ or $4T_s/3$, which is four times of the camera sensor captured time. In other words, one symbol with a length of $4T_s/3$ is captured using four frames. This setting constraint can mitigate the occurrence of frame overlapping. Fig. 8 shows the occurrence of frame overlapping in $-\eta T_c \leq t_d < 0$ (at Frame #1). For $0 \leq t_d \leq (1-\eta)T_c$, there is no frame overlapping. The probability of occurrence of frame overlapping is equal to the exposure time ratio of the camera sensor, $P(\xi) = \eta$. The frame overlapping

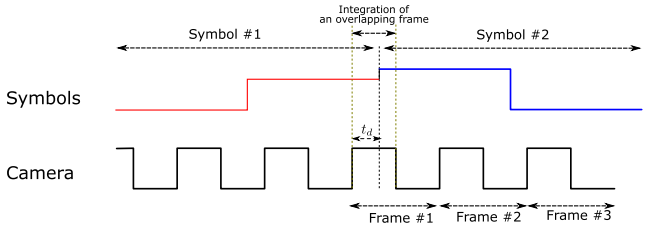


FIGURE 8. Frame overlapping occurred in $-\eta T_c \leq t_d < 0$, where the camera sensor integrates different parts of symbols, Symbol #1 and Symbol #2, into its own frame.

only occurs on the first frame (or the fourth frame, depends on our perspective). Therefore, in order to avoid this problem, a *frame discarding* technique is applied to every first frame no matter whether frame overlapping happens or not. Then, the decoding function starts from the second to the fourth frames. Most importantly, the preamble signal is not only for the calculation of the amplitude and phase of a symbol but also for the first frame detection of other symbols.

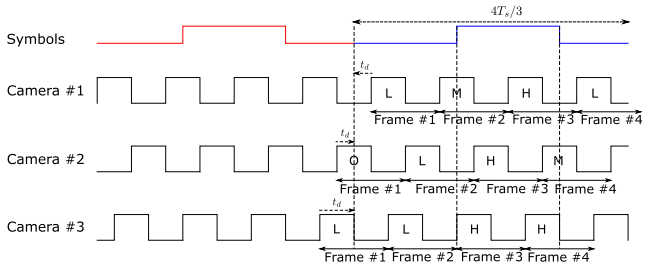


FIGURE 9. Preamble patterns for different t_d in which the *high* frames are always in the third position. (Note: *O*, *L*, *M*, and *H* denote *overlapping*, *low*, *mid*, and *high* frames, respectively.).

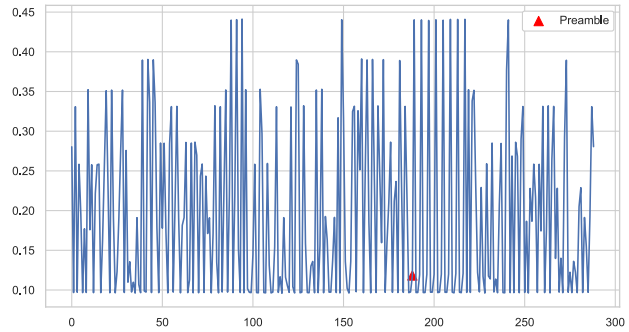
From Fig. 9, there are three possible 4-frame patterns of the preamble signals repeated 8 times as below

1) $L - M - H - L$ pattern for $0 \leq t_d \leq (1-\eta)T_c$,
2) $O - L - H - M$ pattern for $-\eta T_c < t_d < 0$, and
3) $L - L - H - H$ pattern for $t_d = \eta T_c$.

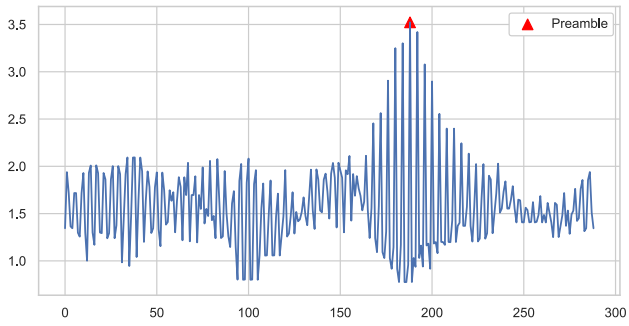
However, the *overlapping* (O) frame, which has *low* (L) value in between two preamble symbols, occurs in the second pattern, where the *mid* frame comes after the *high* frame. By using this pattern, we can easily discard the first frame of the symbols with the rule *"the first high frame in the preamble signals is the third frame of the first preamble symbol."* This leads to an update of the phase shift of preamble symbols as defined in (7). Therefore, the valid frames for decoding start from the second to the fourth frames for each symbol. Then, one can calculate the time difference t_d using (6).

By those three patterns, one can do a trick for automatic preamble detection using a convolution method, because those patterns have at least one H-frame for each preamble symbol that is repeated 8 times. Then, we can construct a convolution kernel $k_8 = [0_1, 1_1, 0_1, 0_1, 0_2, 1_2, 0_2, 0_2, \ldots, 0_8, 1_8, 0_8, 0_8]$ and convolve it with the frame intensity signal. Consequently, the maximum value of the convolution result tells us where the starting point of the preamble symbols

block is. In practice, we do not need to convolve all the frame intensities because we have a specific packet format, i.e., 8 preamble and 64 data symbols. We only convolve the first $(8 + 64 + 8) \times 4 = 320$ frame intensities resulting from 1 block of data symbols and 2 blocks of preamble symbols length containing 4 frames for each. The kernel k_8 detects the H-frame at the second frame so that the first to the third frames are used for the decoding calculation. Fig. 10 shows a real signal obtained by a camera and its convolution to detect the starting point of preamble symbols block.



(a) The frame intensity sequences



(b) The result of convolution of (a) and k_8

FIGURE 10. Preamble detection using convolution kernel k_8.

D. BEYOND $\eta = 0.5$ AND $T_C = T_S/3$ SETTING

In this subsection, we illustrate that different settings of a camera sensor have different impacts on a pair of Tx-Rx. For example, an LED flicker problem can occur at the transmitter and a frame overlapping problem can occur at the receiver. Three important parameters, η, T_c, and T_s, and their configurations can be explained as the following three cases.

1) INCREASING T_C WITH A FIXED $\eta = 0.5$

The period of the camera sensor T_c is varied by k in two subcases: T_{c_1} and T_{c_2} as defined by (9) and (10).

$$T_{c_1} = (2k + 5/3)T_s, \quad \text{for } k = 0, 1, 2, \ldots, n, \tag{9}$$
$$T_{c_2} = (2k + 1/3)T_s, \quad \text{for } k = 0, 1, 2, \ldots, n. \tag{10}$$

Equations (9) and (10) represent camera #2 ($k = 0$) and camera #3 ($k = 1$) settings, respectively, as illustrated in Fig. 11.

The two settings have different lengths, T_{c_1} and T_{c_2}, leading to a different ordering of intensity functions as illustrated in Table 2. The default order of the intensity functions for Frame #1, Frame #2, and Frame #3 is $r_1(t)$, $r_2(t)$, and $r_3(t)$. However, when we apply the different equation settings to the system, this leads to a different ordering of the frame intensity functions. For (9), the order of the intensity functions for each frame is different from the default order, as can be seen in the first row of Table 2, where the intensity functions of Frame #1, Frame #2, and Frame #3 are ordered as $r_2(t)$, $r_1(t)$, and $r_3(t)$. For (10), when $k = 0$, the result setting is the same as the default setting, $T_c = T_s/3$ which is the default order, as can be seen in the third row of Table 2. Both equations show that even though the exposure time ratio is kept at $\eta = 0.5$, the longer T_c setting still provides the correct decoding. This can be proved in Section III-D4. This is a flexibility feature of our proposed scheme. In another point of view, if we fix the frequency of a camera sensor to $f_c = 1/T_c$, and increase the frequency of symbols to $f_s = 1/T_s$ by increasing the value of k, then this setting can mitigate the LED flicker problem at the transmitter.

TABLE 2. The intensity function of each frame based on the equation that is used.

No.	Equation	Frame #1	Frame #2	Frame #3
1	(9)	r_2	r_1	r_3
2	(11)	r_1	r_3	r_2
3	(10) and (12)	r_1	r_2	r_3

2) INCREASING T_C WITH A SETTING $\eta < 0.5$

The period of the camera sensor T_c is varied by k in two subcases: T_{c_3} and T_{c_4} as defined by (11) and (12).

$$T_{c_3} = (k + 2/3)T_s, \quad \text{for } k = 0, 1, 2, \ldots, n, \tag{11}$$
$$T_{c_4} = (k + 1/3)T_s, \quad \text{for } k = 0, 1, 2, \ldots, n. \tag{12}$$

Equations (11) and (12) represent camera #4 ($k = 0$) and camera #5 ($k = 1$) settings, respectively, as illustrated in Fig. 11, and orders of their intensity functions are shown in the second and third rows of Table 2, respectively. Here, we can increase the T_c while fixing the exposure time to

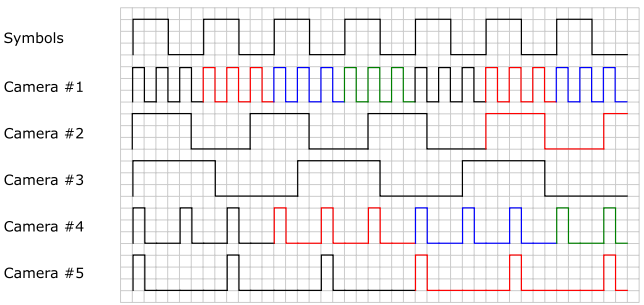


FIGURE 11. Different values of camera period T_c and exposure time ratio η for several proper settings.

$\eta T_c = T_s/6$. Again, in the case of (12) for $k = 0$, we get the same result as that from the default setting. From (11) and (12), the exposure time ratios, η_1 and η_2, can be determined as follows,

$$\eta_1 = \frac{T_s/6}{(k+2/3)T_s} = \frac{1}{2(3k+2)}, \tag{13}$$

$$\eta_2 = \frac{T_s/6}{(k+1/3)T_s} = \frac{1}{2(3k+1)}, \tag{14}$$

respectively. Both (13) and (14) show that the overlapping probability, $P(\xi) = \eta$, can be reduced by setting k to a high value. For instance, if $k = 8$, the overlapping probabilities determined by (13) and (14) are $P(\xi) = 2\%$ and $P(\xi) = 1.9\%$, respectively. These very small overlapping probabilities make it possible to use only three frames for decoding regardless of frame discarding method, thus leading to an increase in bit rate of around 32%. Moreover, we can develop an algorithm for detecting the overlapping occurrence in advance. If overlapping occurs, the camera can be reset with a small jitter time. This technique can solve the overlapping frame problem. Therefore, with this setting, we are not only able to increase the symbol frequency f_s for solving the flicker problem but also to decrease the probability of frame overlapping.

3) GENERAL SETTING OF η AND T_C

As a result of the previous setting in Subsection III-D2, if T_c of (11) and (12) are increased by an integer factor n that is not evenly divisible by 3, and the exposure time is still fixed, we have the following relation.

$$T_c = nT_s/3 \ \text{ for } n \bmod 3 \neq 0 \tag{15}$$

Thus, n can be $3k+2$ or $3k+1$ which is related to (11) or (12), respectively. Let us define $T_{c_{base}}$ to be the capture period of the camera before we increase it by the factor n or when $n = 1$. From the default setting, we have $T_{c_{base}} = T_s/3$ and $\eta_{base} = 0.5$. Thus, substituting this to (15) leads to

$$T_c = nT_{c_{base}} \ \text{ for } n \bmod 3 \neq 0. \tag{16}$$

Here, we can choose any $T_{c_{base}}$ based on (9) and (10). Basically, (16) has $T_{c_{base}}$ obtained from (10) with $k = 0$. However, we can change it by using different settings of k in both equations.

Furthermore, the exposure time of this setting is $\eta_{base}T_{c_{base}}$. Thus, the exposure time ratio is

$$\eta = \frac{T_{c_{base}}\eta_{base}}{nT_{c_{base}}} = \frac{1}{2n} \text{ for } n \bmod 3 \neq 0, \tag{17}$$

which is highly related to (13) and (14) with $n = 3k + 2$ or $n = 3k + 1$, respectively.

Note that all of the settings mentioned so far are called proper settings whenever we can uniquely define the phase and the amplitude of a square wave signal as defined in **Definition 1**. In this case, there is no ambiguous area in the intensity functions of the three frames as shown in Fig. 4(b). For (15), n cannot be in the form of $3k$, because it makes all of the intensity functions of the three frames to be the same, $r_1(t) = r_2(t) = r_3(t)$, because all three exposure times of the frames integrate in the same area of a symbol. However, all of those settings have different formulations of frame intensity such that the intensity functions can be swapped depending on a given setting, as illustrated in Table 2. In this case, if we use $T_{c_{base}}$ from (9), the first and second intensity functions are swapped, otherwise they remain the same. Also, if we use $n = 3k + 2$ setting in (15), the second and third intensity functions are swapped, otherwise they remain the same.

4) THE FRAME INTENSITY FUNCTION

There are 4 equations: (9), (10), (11), and (12) that show the relations between T_s and T_c. However, the intensity of each frame is different from the previous functions (2) and (3). Some of the intensity functions are swapped, especially in (9) and (11). The relation between each frame and its intensity for each equation is shown in Table 2. In this case, the decoding function (4) depends on the conditional table of intensity functions r_1(t), r_2(t) and r_3(t), but the use of a different setting might affect the swapped intensity functions of some frames. Ones should refer to Table 2 that is used for locating in which frame the intensity functions $r_1(t)$, $r_2(t)$, and $r_3(t)$ are. Hence, we can use the phase position i in Table 1.

Now, let us go to the proof of the order of intensity function in Table 2. A proof of these settings means that these proposed settings are proper according to **Definition 1**. From the default setting, the intensity function is set with $\eta = 0.5$ and $T_c = T_s/3$ or the same setting as (10) for $k = 0$. This setting provides three intensity functions in the order of $r_1(t)$, $r_2(t)$, and $r_3(t)$, which are the intensities of the first, second, and third frames, respectively. With this setting, we have an exposure time $\eta T_c = T_s/6$. Thus,

$$\begin{aligned} r_1(t) &= \int_0^{\eta T_c} s(t)dt = \int_0^{T_s/6} s(t)dt, \\ r_2(t) &= \int_{T_s/3}^{T_s/2} s(t)dt = r_1(t - T_s/3), \\ r_3(t) &= \int_{2T_s/3}^{5T_s/6} s(t)dt = r_1(t - 2T_s/3), \end{aligned} \tag{18}$$

where $s(t)$ is the periodical square wave signal with a period of T_s that has the following properties,

$$\begin{aligned} s(t + kT_s) &= s(t) \\ &= \begin{cases} Amplitude & t \in [0, T_s/2] \\ 0 & t \in [T_s/2, T_s] \end{cases}, \\ \int_{kT_s}^{kT_s+c} s(t)dt &= \int_0^c s(t)dt, \end{aligned} \tag{19}$$

where k is an integer value. One period intensity functions are plotted and shown in Fig. 4(b) which is the same as those expressed by (2) and (3). Now, the frame intensity functions

of (9) are defined as follows,

$$\begin{aligned}
\text{Frame}_1 &= \int_0^{\eta T_c} s(t)dt = \int_0^{\eta(2k+1+2/3)T_s} s(t)dt \\
&= \int_0^{(k+1/2+1/3)T_s} s(t)dt \\
&= \int_0^{kT_s} s(t)dt + \int_{kT_s}^{kT_s+T_s/3} s(t)dt \\
&\quad + \int_{kT_s+T_s/3}^{kT_s+5T_s/6} s(t)dt \\
&= \int_0^{kT_s} s(t)dt + \int_0^{T_s/3} s(t)dt + \int_{T_s/3}^{5T_s/6} s(t)dt \\
&= A + B + \int_{T_s/3}^{5T_s/6} s(t)dt \\
&= A + B + \int_{T_s/3}^{T_s/2} s(t)dt + \int_{T_s/2}^{5T_s/6} s(t)dt \\
&= A + B + r_2(t) + 0 \\
&= A + B + r_2(t) \\
\text{Frame}_2 &= \int_{(2k+1+2/3)T_s+\eta T_c}^{(2k+1+2/3)T_s+\eta T_c} s(t)dt \\
&= \int_{2T_s/3}^{2T_s/3+\eta T_c} s(t)dt = \int_0^{\eta T_c} s(t-2T_s/3)dt \\
&= A + B + r_2(t-2T_s/3) \\
&= A + B + r_1(t), \\
\text{Frame}_3 &= \int_{2(2k+1+2/3)T_s+\eta T_c}^{2(2k+1+2/3)T_s+\eta T_c} s(t)dt \\
&= \int_{T_s/3}^{T_s/3+\eta T_c} s(t)dt = \int_0^{\eta T_c} s(t-T_s/3)dt \\
&= A + B + r_2(t-T_s/3) \\
&= A + B + r_3(t). \qquad (20)
\end{aligned}$$

Here, $A + B$ is an additional background intensity that will be removed by using (5). Thus, the intensity functions of the first, second, and third frames are r_2, r_1, and r_3, respectively. Using the same kind of derivation, we can prove the relations between the frame and intensity function for the other settings in Table 2.

In addition, the preamble and overlapping detection might be different as we explained previously due to the swapped frame intensity functions.

E. WIENER DECONVOLUTION

In real implementation, since light propagates to the nearest area around the light source, an LED light point captured by a camera is not a single circle with uniform intensity. This phenomenon has an impact on the usage of multiple LEDs and an image sensor for VLC. Fortunately, an LED light point can be represented by a point spread function (PSF), which is very useful for the case the lights from two or more LEDs interfering with each other. For example, if two LEDs are placed closely to each other, the spread point of one LED might affect that of the other LED in the captured image. In this case, we can apply *Wiener* deconvolution [23] to the captured image together with the PSF from the LED so that we can remove the interference between LEDs. However, we need to estimate the PSF before we can apply *Wiener* deconvolution. The *Wiener* filter is derived as follows. Suppose that we have a system that can be mathematically expressed by

$$g(x, y) = f(x, y) * h(x, y) + n(x, y), \qquad (21)$$

where $f(x, y)$, $h(x, y)$, $n(x, y)$, and $g(x, y)$ are the original image, PSF, noise, and captured image, respectively, and * is the convolution operator. Now, to estimate the original image $f(x, y)$, a *Wiener* filter for deconvolving the captured image can be determined in Fourier domain as below

$$Wiener = \frac{\bar{H}(u, v)}{|H(u, v)|^2 + 1/snr}, \qquad (22)$$

where $\bar{H}$ is the complex conjugate of H and snr is the given signal-to-noise ratio parameter. From this expression, the original image can be estimated by

$$F(u, v) = \frac{\bar{H}(u, v)}{|H(u, v)|^2 + 1/snr} G(u, v), \qquad (23)$$

where $F(u, v)$, $H(u, v)$, and $G(u, v)$ are the Fourier transforms of $f(x, y)$, $h(x, y)$, and $g(x, y)$, respectively. The snr ratio can be a constant value or a matrix.

1) PSF ESTIMATION

Here, we need to estimate the PSF for *Wiener* deconvolution. For simple estimation, we can assume that the PSF is the same as a point spread of a single LED. Thus, we can estimate the PSF by capturing a single LED point image. This might not be the best estimation of PSF because the captured image can have a noise. However, it works very well in our case as shown in Fig. 12 that separates the LED points more distinctly. For a more advanced method, one can model the PSF so that it can approximate the spreads of the LED point. However, in our work, we just focus on the proof of concept for SW-QAM using multiple LEDs, in this case, PSF estimation using LED point image is sufficient for the investigation.

(a) Image (b) PSF (c) Result

FIGURE 12. Applying *Wiener* filter to (a) the captured image with (b) the PSF image and the result (c) that separates the LED points more distinctly.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the performance of the proposed square wave quadrature amplitude modulation (SW-QAM) scheme for

visible light communication (VLC) was evaluated in terms of error rate E_r defined by (24).

$$E_r = 100\% \times N_{inval}/N_T, \tag{24}$$

where N_{inval} and N_T denote the number of invalidly reconstructed symbols and the number of totally valid symbols, respectively. In other words, the error rate can be viewed as a misclassified symbol rate or the ratio between the number of misclassified symbols to the number of the entire set of symbols. We tested our SW-QAM in different settings, i.e., LED frequencies, distances, and exposure time ratios. We also tested the SW-QAM on multiple LEDs scheme.

A. EQUIPMENT SETUP

At the transmitter side, LEDs (OptoSupply OSB56A5111A) were used as the transmitter. A dimming paper was placed in front of the LEDs to soften their light. The function generator (NF-WF1948) was used for generating the encoded signals. A 64-QAM constellation scheme was used with preamble symbols encoded by setting $\theta = \pi$ and $A = 1$ as depicted in Fig. 13. Eight preamble symbols and 64 symbols were encoded by the function generator and sent the encoded signal repeatedly as LED intensity. At the receiver side, the global shutter camera (Point Grey FL3-U3-13Y3M) was used as the receiver since we tried to conceptually prove our SW-QAM scheme. It captured 2,000 sequential images with a resolution of 128×128 pixels and a frame rate of 60 *fps* ($f_c = 60$ Hz) for each run. The automatic preamble detection and frame discarding were applied for all experiments as described in Section III-C since the camera was not synchronized with the LEDs. In addition to the captured image, we took the average intensity of a 13×13 pixel square sampling block at the center of each LED and used it to estimate the current frame intensity.

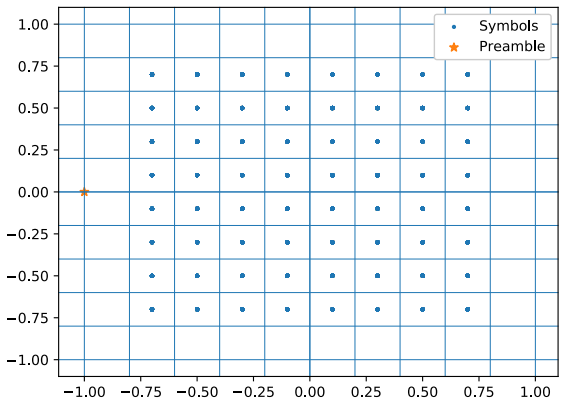


FIGURE 13. A 64-QAM constellation of the symbols and the preamble symbols at phase $\theta = \pi$ and amplitude $A = 1$.

B. EXPERIMENTS WITH A SINGLE LED SETTING

We tested our SW-QAM scheme with a single LED transmitter on several settings, i.e., two symbol frequencies f_s, two exposure time ratios η, and two lighting conditions, i.e., dark (turn off all of the lamps except computer screen) and bright (turn on all of the lamps). The experimental room is $3 \times 3\ m^2$ with four room lightings. Further, all of above settings were tested with three different Tx-Rx distances: $d = 30$ cm, $d = 50$ cm, and $d = 100$ cm as schematically shown in Fig. 14.

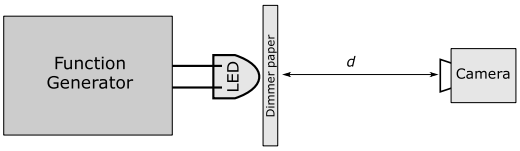


FIGURE 14. Experimental setup for single LED setting.

1) $F_S = 20$ HZ WITH $\eta = 0.5$

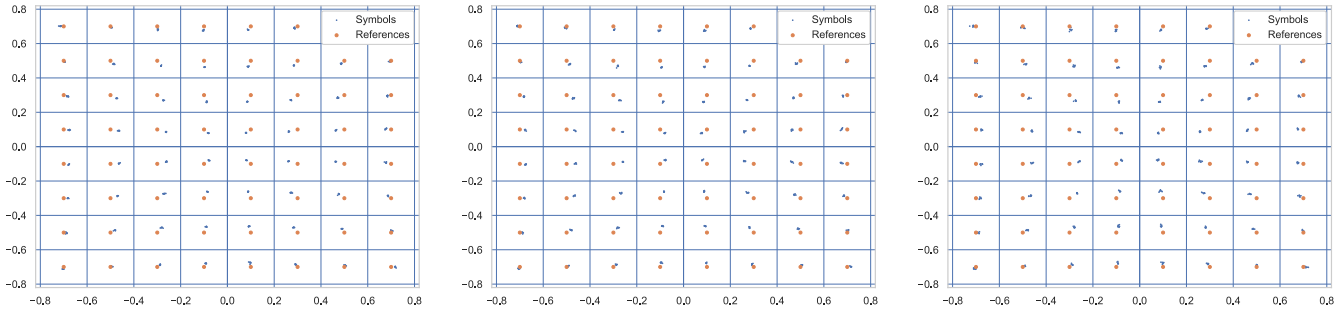
Here, we set symbol frequency $f_s = 20$ Hz and exposure time ratio $\eta = 0.5$, i.e., Eq. (10) with $k = 0$ (default setting). The camera aperture was set to $f/6$. Fig. 15 shows the 64-QAM constellation from the received signal for both lighting conditions, dark and bright. Clearly, the SW-QAM decoding method was able to decode all symbols without any errors – see Table. 3. Further, even in hard conditions, i.e., far distance, 100 cm, and bright condition (Fig. 15(f)), our SW-QAM still performed very well. The far distance leads to the degradation of signal strength while the bright lighting condition gives more noise than the dark lighting condition. Thus, it will reduce the decoding performance.

2) $F_S = 140$ HZ WITH $\eta = 0.5$

By increasing f_s greater than 100 Hz, we can mitigate the LED flicker problem. Therefore, here we set the symbol frequency $f_s = 140$ Hz and exposure time ratio $\eta = 0.5$, i.e., Eq. (10) with $k = 7$. The camera aperture was set to $f/6$. For the dark lighting condition, this setting performed well, without any errors, although the spread of points (decoded symbols) was as not good as that of 20 Hz setting, as depicted in Fig. 16. The decoded symbols spread far from the reference point but they still inside the boundary (a square border), so in this case, there is no error symbol – see Table. 3. On the other hand, for the bright lighting condition, the performance of SW-QAM was degraded, since the decoded symbols spread far from their reference points, as can be seen in Table. 3. If the camera

TABLE 3. Symbol error rate, resulting from the different settings, different lighting conditions, and different distances, of a single LED point transmitter.

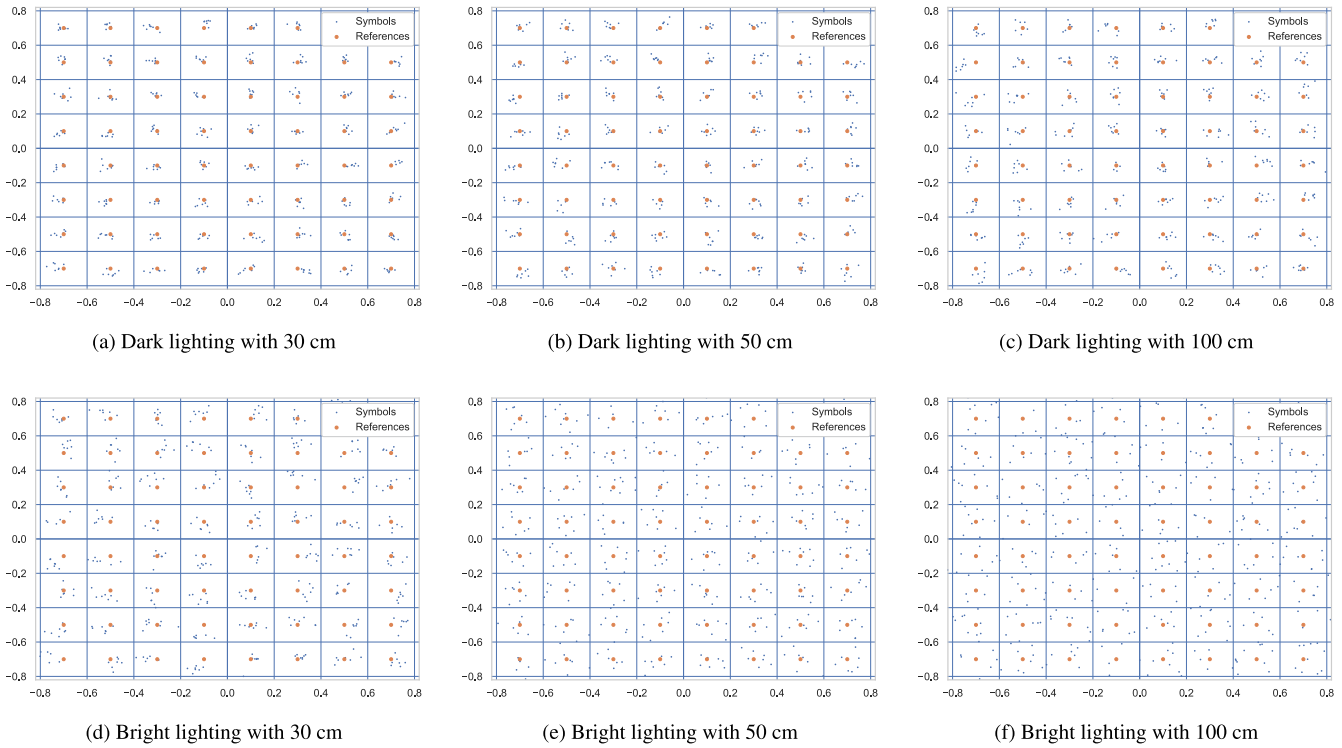
Setting	Lighting condition	Symbol error rate (%) 30 cm	50 cm	100 cm
20 Hz, $\eta = 0.5$, $f/6$, Eq. (10) with k=0	Dark	0	0	0
	Bright	0	0	0
140 Hz, $\eta = 0.5$, $f/6$, Eq. (10) with k=7	Dark	0	0	0
	Bright	0.45	2.26	8.78
140 Hz, $\eta = 1/14$, $f/2$, Eq. (12) with k=7	Dark	0	0	0
	Bright	0	0	0



(a) Dark lighting with 30 cm (b) Dark lighting with 50 cm (c) Dark lighting with 100 cm

(d) Bright lighting with 30 cm (e) Bright lighting with 50 cm (f) Bright lighting with 100 cm

FIGURE 15. The performance of SW-QAM when setting $f_S = 20$ Hz and $\eta = 0.5$ for all test cases and varying different settings of lighting conditions and distances.

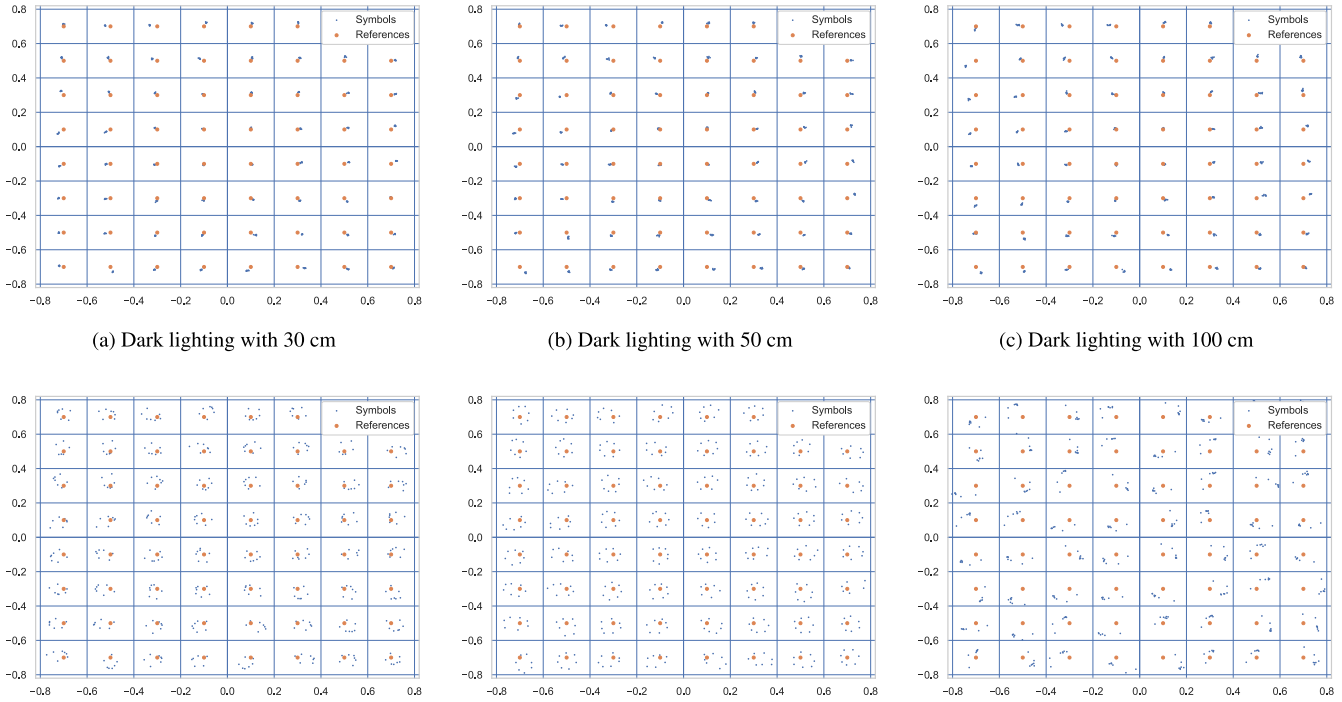


(a) Dark lighting with 30 cm (b) Dark lighting with 50 cm (c) Dark lighting with 100 cm

(d) Bright lighting with 30 cm (e) Bright lighting with 50 cm (f) Bright lighting with 100 cm

FIGURE 16. The performance of SW-QAM when setting $f_S = 140$ Hz and $\eta = 0.5$ for all test cases and varying different settings of lighting conditions and distances.

is far from the LED, the error rate is increasing as can be seen in Fig. 16, which shows that the decoded symbols spread far from the reference point. However, the error rate is still below 10%. Here, the 20 Hz setting performed better than the 140 Hz setting because r_{max} of 140 Hz setting was lower than that of 20 Hz due to the increasing background intensity.



(a) Dark lighting with 30 cm (b) Dark lighting with 50 cm (c) Dark lighting with 100 cm

(d) Bright lighting with 30 cm (e) Bright lighting with 50 cm (f) Bright lighting with 100 cm

FIGURE 17. The performance of SW-QAM when setting $f_S = 140$ Hz and $\eta = 1/14$ for all test cases and varying different settings of lighting conditions and distances.

In other words, the 140 Hz setting had a lower signal strength than 20 Hz setting, with the same exposure time ratio η and camera aperture.

3) $F_S = 140$ HZ WITH $\eta = 1/14$

For another LED flicker-free setting, we set $f_s = 140$ Hz and $\eta = 1/14$, i.e., Eq. (12) with $k = 7$. Here, because the exposure time was small, so we increased the amount of light coming to the image sensor by setting the camera aperture to $f/2$ so that we had a good strength of the received signal. Fig. 17 shows the 64-QAM constellation from the received signal for both lighting conditions, dark and bright. For the dark condition, this setting results in small spreads which are much better than the 140 Hz and $\eta = 0.5$ setting. It resulted zero error rate – see Table. 3. For the bright condition, the decoded symbol spread becomes larger when the distance between Tx and Rx is farther. However, the SW-QAM still performed well, without any errors. We can say that the setting of the larger camera aperture can compensate for the longer distance between Tx and Rx. However, for the previous $f_s = 140$ Hz with $\eta = 0.5$ setting, the aperture was small, because the large aperture will lead to overexposure pixels or white pixels, i.e., information loss. In addition, the advantage of the small exposure time ratio leads to a small frame overlapping probability. Thus, we can use the 3-frame length of symbols for the SW-QAM to increase the communication speed when subjected to that frame overlapping probability.

In summary, we can use different symbol frequencies, f_s, to mitigate the flicker problem. On the other hand, the camera frequency, f_c, only affects the communication speed because the camera is used as a sampling machine, the higher the camera frequency, the faster the speed is.

C. EXPERIMENTS WITH MULTIPLE LED SETTING

For multiple LED setting, at the transmitter side, two LEDs were synchronized and installed very close to each other, only 2 millimeters apart as schematically shown in Fig. 18. With this setting, light interference is present. The symbol frequency f_s was sent at 20 Hz generated from a function generator. At the receiver side, the global shutter camera was placed 30 cm far from the LEDs. The camera configuration was set with a frame rate of $f_c = 60$ Hz, an exposure time ratio of $\eta = 0.5$ and an aperture of $f/6$. Five runs of the same experiment were conducted in a dark lighting condition to strengthen the result. Thus, we had a total of $5 \times 2,000$ images in our collection, i.e., five different preamble positions and noises. The *snr* parameter of *Wiener* deconvolution was set

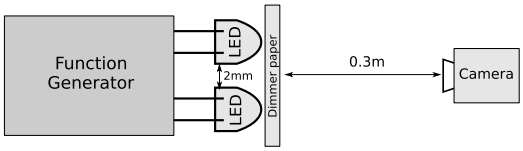
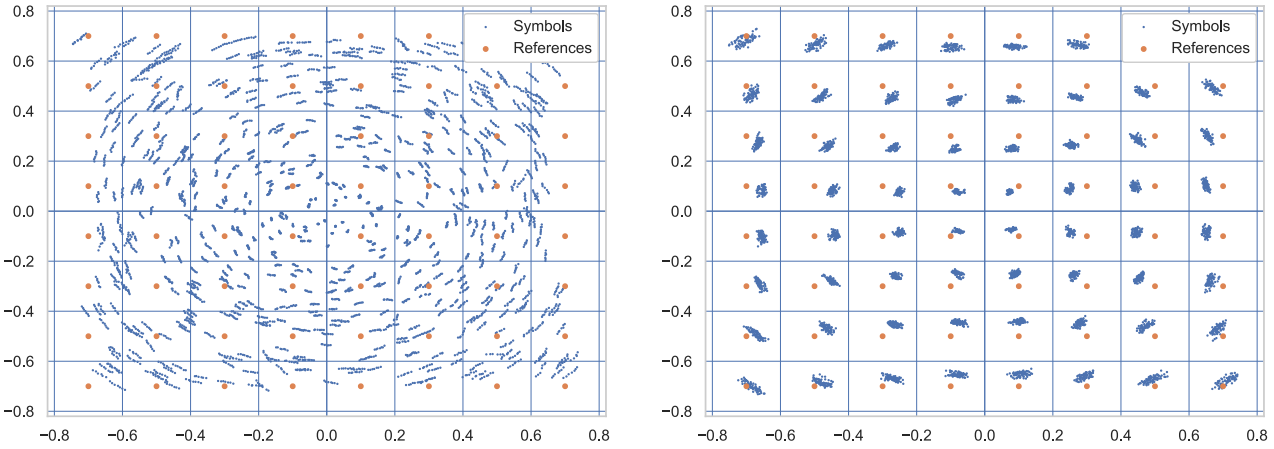


FIGURE 18. Experimental setup for multiple LED setting.



(a) SW-QAM

(b) SW-QAM with $Wiener$ filter

FIGURE 19. Performances of SW-QAM with and without the use of *Wiener* filter in 64-QAM constellation.

to a constant value of 0.02 for simplification. For the PSF, we estimated using a single predefined LED image used for all frames.

Here, we compared the performance of our schemes, i.e., ordinary SW-QAM and SW-QAM with *Wiener* filter. The comparative performance of both proposed methods in terms of error rate is shown in Table 4. It can be seen that the performance of the ordinary SW-QAM method is much lower than that of the SW-QAM with *Wiener* filter; that is, the error rate is 50.02% for the ordinary one while it is 0.00%, without any errors for the improved one. These results showed that, as the two LEDs were installed very close to each other, light interference tended to occur. The SW-QAM with *Wiener* filter method was able to correctly decode better than the ordinary SW-QAM method. This achievement came from the use of *Wiener* filter.

TABLE 4. An error rate comparison of both proposed methods.

Method	Error rate (%)
SW-QAM	50.02
SW-QAM with $Wiener$ filter	0.00

Let us look at the performances of both schemes, ordinary SW-QAM and SW-QAM with *Wiener* filter, more closely, as shown in Fig. 19. The plot shows how light interference affects the decoded symbols in the SW-QAM method, while the *Wiener* filter is able to reduce the effect of light interference and dramatically improve the performance. For a better result, we need a specific type of PSF for this filter. In our paper, we captured a single LED image as a PSF. An alternative way to reduce the effect of light interference is to separate the LEDs far apart from each other, instead of the short 2 mm distance.

D. DISCUSSION

From the scheme proposed in Section III to its implementation, the experimental results clearly show that SW-QAM is practical and works well in real situations. The proposed SW-QAM method also provides flexible settings in the adjustment of the encoding and decoding parameters.

1) The high frequency setting of symbols/LED, f_s, can solve a flicker problem as reported in [11], [16].
2) A decrease in exposure time ratio η can also decrease the probability of frame overlapping occurrences. With a very low η, the high confidentiality $1-\eta$ of frame overlapping does not occur. Therefore, we can use only three frames for decoding the encoded symbols. However, the proper settings as expressed in (9), (10), (11), and (12) must be satisfied in every situation.
3) We can synchronize the camera and LED signals by using the phase shifting of preamble symbols, t_d, with the same technique described in [16], by shifting the camera signal by t_d. This technique provides a valid decoding function without the need to use a frame discarding method. In other words, three frames are sufficient for decoding. However, in this work, we did not synchronize the camera, thus we used 4-frame symbols.
4) Communication speed can be increased by using a high speed camera (higher f_c) or by increasing the number of LEDs. However, light interference may cause unreliable decoding. A solution is to use the *Wiener* filter or to separate the LEDs far apart from each other or to use different types of LEDs that have a small interference propagation, e.g., flat LEDs.
5) To make the communication channel more reliable, one can reduce the number of bits per symbol by using a smaller QAM encoding diagram such as 16-QAM that has four bits per symbol. The use of this flexible setting depends on the type of application and the condition of the communication channel.

Furthermore, the effective communication speed *bps* depends on several parameters, i.e., the number of LEDs *n*, the number of bit per symbol *bpsym*, the camera speed *fps*,

the preamble ratio pr, and the number of frames per symbol $fpsym$. From these parameters, we can formulate the communication speed as follows.

$$bps = n \times bpsym \times fps \times (1 - pr)/fpsym. \quad (25)$$

In our experiment, the number of LEDs, $n = 1$; the number of bits per symbol, $bpsym = 6$ (64-QAM); the camera speed, $fps = 60$; the preamble ratio, $pr = 8/(8+64) = 0.11$; and the number of frames per symbol, $fpsym = 4$. Hence, the communication speed is $1 \times 6 \times 60 \times (1 - 0.11)/4 = 80.1$ bps. Furthermore, for two LEDs, $n = 2$, the communication speed is 160.2 bps.

Based on (25), our proposed scheme is remarkably easy to scale up the communication speed although a low sampling rate camera, global shutter camera, is used. In our experimental result, we reach 160.2 bps which is a very high communication speed when compared to those that reported in [6], [19], [21]. It can be confirmed that the use of the low speed camera or low sampling rate camera is able to achieve a very high communication speed. Moreover, if an LED array is installed at the transmitter side and a high speed camera is used at the receiver side, then much higher communication speed will be achieved. We can say that the proposed scheme provides scalability of communication speed.

V. CONCLUSION

In this paper, we have proposed a novel modulation method based on the square wave signal called a square wave quadrature amplitude modulation (SW-QAM) method for visible light communication (VLC). The proposed SW-QAM method can be expressed by generic and simple mathematics. In practice, the transmitter side consists of an SW-QAM encoder and LEDs. The SW-QAM encoder can work with multiple LEDs to increase the bit rate of communication and it also can support a flicker-free modulation scheme. On the other hand, the receiver side consists of an SW-QAM decoder and an image sensor. The SW-QAM decoder can effectively respond to different proper settings of the camera in terms of exposure time ratio. Based on our experimental setup and data collection, the results show that the SW-QAM decoder can decode the encoded signal with a very low error rate in hard condition, i.e., far distance and bright lighting condition. The main achievements of the proposed method are (i) the method is able to provide a scalable communication speed from medium to high depending on the number of LEDs installed at the transmitter side and (ii) the method is able to support the flicker-free modulation scheme by increasing the symbol sending frequency.

For future work, the proposed SW-QAM method still has a room for performance improvement, especially in the part of the SW-QAM decoder that requires complementary techniques such as *Wiener* deconvolution and LED localization. One of the feasible solutions for performance improvement is the development of a fully automatic SW-QAM decoder.

ACKNOWLEDGMENT

The authors would like to thank their colleagues–Mr. Shogo Amaya and Mr. Shota Shimada–who have helped us to collect all of the datasets and provided a very useful discussion on this work. They would also like to thank Hokkaido University for providing the equipment for the experiment.

REFERENCES

[1] M. Figueiredo, L. N. Alves, and C. Ribeiro, "Lighting the wireless world: The promise and challenges of visible light communication," *IEEE Consum. Electron. Mag.*, vol. 6, no. 4, pp. 28–37, Oct. 2017.

[2] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: A survey, potential and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2047–2077, 4th Quart., 2015.

[3] J. Rufo, J. Rabadan, F. Delgado, C. Quintana, and R. Perez-Jimenez, "Experimental evaluation of video transmission through LED illumination devices," *IEEE Trans. Consum. Electron.*, vol. 56, no. 3, pp. 1411–1416, Aug. 2010.

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

[5] F. Wu, L. Chen, S. Cai, and W. Wang, "Experimental study and application of response mask invariant characteristic for generalized visible light communication channel," *IEEE Photon. J.*, vol. 10, no. 3, Jun. 2018, Art. no. 7903912.

[6] T. Nguyen, A. Islam, and Y. M. Jang, "Region-of-interest signaling vehicular system using optical camera communications," *IEEE Photon. J.*, vol. 9, no. 1, Feb. 2017, Art. no. 7900720.

[7] P. Luo, Z. Ghassemlooy, H. Le Minh, X. Tang, and H.-M. Tsai, "Undersampled phase shift ON-OFF keying for camera communication," in *Proc. IEEE 6th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2014, pp. 1–6.

[8] R. D. Roberts, "Undersampled frequency shift ON-OFF keying (UFSOOK) for camera communications (CamCom)," in *Proc. 22nd Wireless Opt. Commun. Conf.*, Chongqing, China, May 2013, pp. 645–648.

[9] 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*, Dec. 2012, pp. 1244–1248.

[10] C.-W. Chow, C.-Y. Chen, and S.-H. Chen, "Enhancement of signal performance in LED visible light communications using mobile phone camera," *IEEE Photon. J.*, vol. 7, no. 5, Oct. 2015, Art. no. 7903607. [Online]. Available: http://ieeexplore.ieee.org/document/7265146/

[11] J. D. Bullough, K. S. Hickcox, T. R. Klein, and N. Narendran, "Effects of flicker characteristics from solid-state lighting on detection, acceptability and comfort," *Lighting Res. Technol.*, vol. 43, no. 3, pp. 337–348, 2011.

[12] T. Nagura, T. Yamazato, M. Katayama, T. Yendo, T. Fujii, and H. Okada, "Improved decoding methods of visible light communication system for its using led array and high-speed camera," in *Proc. IEEE 71st Veh. Technol. Conf.*, May 2010, pp. 1–5.

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

[14] T. Nguyen, M. A. Hossain, and Y. M. Jang, "Design and implementation of a novel compatible encoding scheme in the time domain for image sensor communication," *Sensors*, vol. 16, no. 5, p. 736, 2016.

[15] H. Aoyama and M. Oshima, "Visible light communication using a conventional image sensor," in *Proc. IEEE 12th Annu. Consumer Commun. Netw. Conf. (CCNC)*, Jan. 2015, pp. 103–108.

[16] M. Sugimoto, H. Kumaki, T. Akiyama, and H. Hashizume, "Optimally modulated illumination for rapid and accurate time synchronization," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 505–516, Jan. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7572924/

[17] 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 Commun. Mag.*, vol. 52, no. 7, pp. 88–97, Jul. 2014.

[18] Y. Goto, I. Takai, T. Yamazato, H. Okada, T. Fujii, S. Kawahito, S. Arai, T. Yendo, and K. Kamakura, "A new automotive VLC system using optical communication image sensor," *IEEE Photon. J.*, vol. 8, no. 3, Jun. 2016, Art. no. 6802716.
[19] P. Ji, H.-M. Tsai, C. Wang, and F. Liu, "Vehicular visible light communications with LED taillight and rolling shutter camera," in *Proc. IEEE 79th Veh. Technol. Conf. (VTC Spring)*, May 2014, pp. 1–6.
[20] N. Rajagopal, P. Lazik, and A. Rowe, "Visual light landmarks for mobile devices," in *Proc. IEEE 13th Int. Symp. Inf. Process. Sensor Netw.*, Apr. 2014, pp. 249–260.
[21] C. H. Hong, T. Nguyen, N. T. Le, and Y. M. Jang, "Modulation and coding scheme (MCS) for indoor image sensor communication system," *Wireless Pers. Commun.*, vol. 93, no. 4, pp. 987–1003, Apr. 2017. doi: 10.1007/s11277-017-3977-x.
[22] Sony. (2018). *Sony Global—Sony Develops a Back-Illuminated CMOS Image Sensor with Pixel-Parallel A/D Converter That Enables Global Shutter Function*. [Online]. Available: https://www.sony.net/SonyInfo/News/Press/201802/18-018E/index.html
[23] C. Rafael Gonzalez, E. R. Woods, and L. S. Eddins, *Digital Image Processing Using MATLAB*, vol. 10. Upper Saddle River, NJ, USA: Prentice-Hall, 2003.

SYUKRON ABU ISHAQ ALFAROZI (M'18) received the B.Eng. degree in electrical engineering and information technology from the Universitas Gadjah Mada, Yogyakarta, Indonesia, in 2014. He is currently pursuing the Ph.D. degree in information technology with the King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, under the sandwich program with the Hokkaido University, Sapporo, Japan. His research interests include machine learning/deep learning and its application, computer vision, and signal processing.

KITSUCHART PASUPA (M'12–SM'16) received the B.Eng. degree in electrical engineering from the Sirindhorn International Institute of Technology, Thammasat University, Thailand, in 2003, and the M.Sc. (Eng.) and Ph.D. degrees in automatic control and systems engineering from the Department of Automatic Control and Systems Engineering, The University of Sheffield, in 2004 and 2008, respectively. He was a Research Fellow with the University of Southampton and The University of Sheffield. He is currently an Associate Professor with the Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. His main research interest includes the application of machine learning techniques in the real-world application.

HIROMICHI HASHIZUME (M'05) received the B.E., M.E., and D.E. degrees in electronic engineering from The University of Tokyo, Tokyo, Japan, in 1979, 1981, and 1984, respectively. He is currently a Professor with the Information Systems Architecture Science Research Division, National Institute of Informatics, Japan. His research field was computer networks and telecommunication. Recently, his interest has shifted to the mathematical modeling of communication systems.

KUNTPONG WORARATPANYA (M'17) received the B.Ind.Tech. degree in computer technology, the M.Eng. degree in computer engineering, and the D.Eng. degree in electrical engineering from the King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, in 1992, 1996, and 2005, respectively. He is currently an Assistant Professor with the Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang. His research interests include stereoscopic acquisition and compression, multi-media coding and processing, signal processing, speech recognition and processing, pattern recognition and image processing, computer vision, and machine learning/deep learning.

MASANORI SUGIMOTO (M'97) received the B.E., M.E., and D.E. degrees in aeronautics and astronautics from The University of Tokyo, Tokyo, Japan, in 1990, 1992, and 1995, respectively. He is currently a Professor with the Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Japan. His research interests include acoustic engineering, signal processing, artificial intelligence, and human–computer interaction technologies for designing smart systems and environments.