HIGH-SPEED OPTICAL CAMERA COMMUNICATION USING AN OPTIMALLY MODULATED SIGNAL

Masanori Sugimoto[†*] *Hayato Kumaki* [†] *Takayuki Akiyama*[⋆] *Hiromichi Hashizume*[‡]

[†] Hokkaido University [⋆]SOKENDAI [‡]National Institute of Informatics

ABSTRACT

This paper describes a high-speed optical camera communication (OCC) technique using an LED and a rolling-shutter camera. In the proposed technique, the symbols being transmitted are encoded as time delays of optimally modulated signals derived theoretically. A receiver decodes the symbols by using intensities obtained from four consecutive line sensors of a camera. Experiments using a camera having performance similar to that of a general-purpose camera show that the proposed technique can achieve 0.833 $\sim$ 1.17 bits per line sensing and that symbol transmission is possible with a longer exposure time setting; this is difficult to achieve using existing on–off keying OCC techniques.

Index Terms— optical camera communication, optimally modulated signal, exposure time, rolling shutter

1. INTRODUCTION

Because of the huge penetration of smartphones having built-in cameras, the standardization of optical camera communication (OCC) [1] has been investigated recently. Visible light communication (VLC) [2] including OCC cannot always achieve high-speed communication like Wi-Fi but can restrict its communication to line-of-sight areas. VLC is a promising technology for places where Wi-Fi is prohibited, such as hospitals or nuclear power plants. There are many VLC systems for optical cameras, for example, using displays as transmitters ([3], [4], [5]), indoor positioning applications ([6], [7], [8], [9], [10], [11]) and so on.

We previously devised a rapid and accurate time-difference estimation method that uses LED illumination and a general-purpose camera with a global shutter [12]. The paper extends this method of OCC using a rolling-shutter camera such as a smartphone camera. The proposed method uses intensity values of an LED illumination measured by four consecutive line sensors to decode one symbol. For a 60 fps camera with 1,000 line sensors per frame, the fundamental frequency of modulated light from an LED is 15 kHz, and thus human eyes do not perceive the flicker [13]. Many OCC systems using a rolling-shutter camera employ on–off keying to encode symbols [14], [15]. To identify the camera's on–off status correctly, it is not allowed to overlap exposure times of neighboring line sensors. Thus, the exposure time of each line sensor must be very short, which makes it difficult to implement augmented reality applications by superimposing received data through OCC on a captured image. On the other hand, the proposed method encodes symbols based on phase-shift modulation by changing time delays of optimally modulated signals. Therefore, it can increase the exposure times of line sensors.

Evaluations of the proposed method were conducted using a camera with performance similar to that of a general-purpose camera. The results show that it could achieve 0.833 $\sim$ 1.17 bits per line sensor without errors, which is a comparable or better performance than on–off keying-based methods that cannot theoretically exceed 1 bit per line sensor.

The contributions of this paper are as follows:

- We propose an OCC technique based on phase-shift modulation that can achieve high-speed communication.
- We investigate how the proposed method changes its performance for different exposure time settings.

2. PROPOSED METHOD

2.1. Time-Difference Estimation for a Rolling-Shutter Camera

Suppose T_C, T_E, $\eta\ (= T_E/T_C)$, and L are the frame period, exposure time, exposure time ratio, and number of line sensors of a camera, respectively (Figure 1). The period of a signal $s(t)$ emitted from an LED is given as mT_C $(m > 0)$. The time difference δmT_C $(0 \leq \delta < 1)$ is defined as that between a rising edge or peak value time of $s(t)$ and the shutter release time of the first line sensor. The time difference τ_l at a line sensor l $(l = 0, 1, \cdots, L-1)$ is represented as $\tau_l = \delta mT_C + \frac{l}{L}T_C$. The intensity $r(\tau_l)$ obtained by the line sensor is calculated as follows:

$$r(\tau_l) = \frac{1}{T_C}\int_0^{\eta T_C} s(t + \delta mT_C + \frac{l}{L}T_C)dt. \qquad (1)$$

*The work was supported by JSPS KAKENHI under Grants 16H03070 and 17K19952.

The Fourier transform of $r(\tau_l)$ on δ is given as $R_k(l)$ $(k = 0, \pm1, \pm2, \cdots)$:

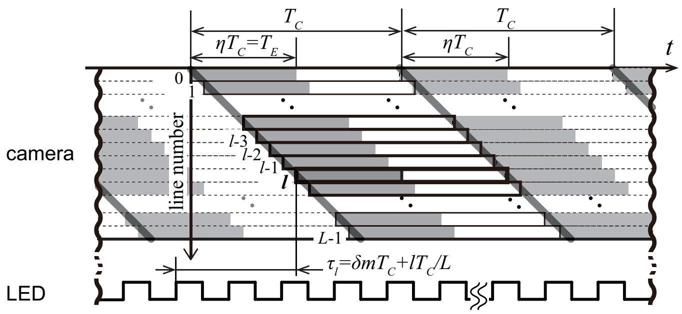

Fig. 1. Exposure timings of line sensors in a rolling-shutter camera.

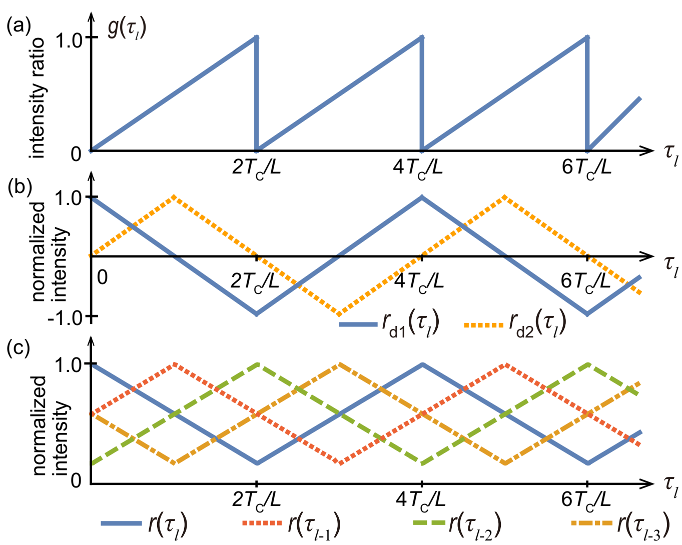

Fig. 2. (a) Intensity ratio function; (b) Intensity difference functions; (c) Intensity functions.

$$\begin{aligned} R_k(l) &= \frac{1}{T_C}\int_0^1\!\!\int_0^{\eta T_C} s(t+\delta m T_C+\frac{l}{L}T_C)dt\, e^{-j2\pi k\delta}d\delta \\ &= S_k \eta e^{j\frac{\pi k\eta}{m}}\,\mathrm{sinc}\left(\frac{\pi k\eta}{m}\right) e^{j\frac{2\pi k l}{mL}}. \end{aligned} \tag{2}$$

Note that S_k is the Fourier transform of $s(t)$. Equations (3) and (4) define the intensity difference and intensity ratio functions, respectively:

$$\begin{aligned} r_{d1}(\tau_l) &= r(\tau_l) - r(\tau_{l-2}) \\ r_{d2}(\tau_l) &= r(\tau_{l-1}) - r(\tau_{l-3}) \end{aligned} \tag{3}$$

$$g(\tau_l) = \begin{cases} \dfrac{r_{d2}(\tau_l)}{2\,(r_{d1}(\tau_l) + r_{d2}(\tau_l))} & (r_{d1}(\tau_l)\, r_{d2}(\tau_l) \geq 0) \\ \dfrac{r_{d1}(\tau_l)}{2\,(r_{d1}(\tau_l) - r_{d2}(\tau_l))} + \dfrac{1}{2} & (r_{d1}(\tau_l)\, r_{d2}(\tau_l) < 0)\,. \end{cases} \tag{4}$$

If $g(\tau_l)$ is obtained as a sawtooth wave as shown in Figure 2(a), the time difference τ_l is proportional to $g(\tau_l)$ during the interval between $2nT_C/L$ and $2(n+1)T_C/L$ (n: integer). This requires that $r_{d1}(\tau_l)$ and $r_{d2}(\tau_l)$ are triangular waves with a phase difference of $\pi/2$ (time difference T_C/L), as shown in Figure 2(b). Figure 2(c) shows that $r(\tau_{l-i})$ $(i = 0, 1, 2, 3)$ must be triangular waves except for a DC component, and their period and phase difference are $4T_C/L$ and $\pi/2$, respectively. The Fourier transform of $r(\tau_l)$ is given as:

$$F_k(l) = \frac{\mathrm{sinc}^2(\frac{k\pi}{2})}{2}\, e^{\frac{jk\pi}{2}l}. \tag{5}$$

When $R_k(l) = F_k(l)$ holds, $s_{(t)}$ is defined as an optimally modulated signal and is represented as $s_{opt}(t)$:

$$s_{opt}(t) = \sum_{k=-\infty}^{\infty} \frac{1}{\eta} e^{-j\frac{k\pi\eta}{m}} \frac{\mathrm{sinc}^2(\frac{k\pi}{2})}{\mathrm{sinc}(\frac{k\pi\eta}{m})} e^{\frac{j\pi k l}{2}(1-\frac{4}{mL})} e^{j\frac{2k\pi}{m T_C}t}. \tag{6}$$

To make $s_{opt}(t)$ independent of l, $m = \frac{4}{L}$ must hold. The value of $\mathrm{sinc}^2(k\pi/2)/\mathrm{sinc}(k\pi\eta/m)$ in (6) must be finite so that $s_{opt}(t)$ always exists, which results in the following equation:

$$\eta = \frac{m}{2p} = \frac{2}{pL} \quad (p = 1, 2, \cdots). \tag{7}$$

$s_{opt}(t)$ is given as (8) $(n = 0, 1, 2, \cdots)$ and its period is $4T_C/L$. Note that $\lfloor x \rfloor$ is the maximum integer not larger than x. Figure 3 shows examples of $s_{opt}(t)$:

$$s_{opt}(t) = \begin{cases} 1 - \dfrac{\eta L}{2}\left\lfloor \dfrac{t - 4nT_C/L}{\eta\, T_C} \right\rfloor & \left(\frac{4nT_C}{L} \leq t < \frac{(4n+2)T_C}{L}\right) \\ \dfrac{\eta L}{2}\left\lfloor \dfrac{t - (4n+2)T_C/L}{\eta\, T_C} \right\rfloor & \left(\frac{(4n+2)T_C}{L} \leq t < \frac{4(n+1)T_C}{L}\right). \end{cases} \tag{8}$$

From Figure 2 (a) and (b), τ_l is easily obtained as the interval between $4nT_C/L$ and $4(n+1)T_C/L$, as shown in Equation (9):

$$\tau_l = \begin{cases} (2\,T_C/L)\; g(\tau_l) & ((r_{d2}(\tau_l) = 0 \text{ and } r_{d1}(\tau_l) > 0) \text{ or } r_{d2}(\tau_l) > 0) \\ (2\,T_C/L)\; (1 + g(\tau_l)) & ((r_{d2}(\tau_l) = 0 \text{ and } r_{d1}(\tau_l) < 0) \text{ or } r_{d2}(\tau_l) < 0). \end{cases} \tag{9}$$

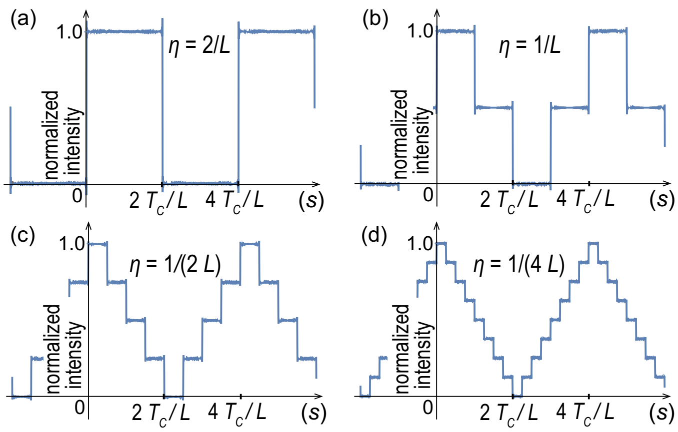

Fig. 3. Optimally modulated signals $s_{opt}(t)$: (a) $\eta = 2/L$, (b) $\eta = 1/L$, (c) $\eta = 1/(2L)$, and (d) $\eta = 1/(4L)$.

2.2. Encoding and Decoding using Time Delay

A symbol given as an N-bit integer $(0 \leq k \leq 2^N - 1)$ is encoded as a time delay t_k $(0 < t_k < 4T_C/L)$ using the following equation at a transmitter:

$$t_k = \frac{2T_C}{L}\left(\frac{k}{2^{N-1}} + \frac{1}{2^N}\right). \tag{10}$$

The time delay $t_k^{'}$ is obtained using Equations (4) and (9) at a receiver. Then, the time delay t_k given at the transmitter is calculated as $t_k^{''}$ using Equation (11). Note that the time difference τ_l between an LED and a line sensor l is estimated before receiving the encoded symbols:

$$t_k^{''} = \begin{cases} t_k^{'} - \tau_l & (t_k^{'} - \tau_l \geq 0) \\ t_k^{'} - \tau_l + \dfrac{4T_C}{L} & (t_k^{'} - \tau_l < 0). \end{cases} \tag{11}$$

The transmitted symbol k is decoded as $k^{''}$ from the time delay $t_k^{''}$ using Equation (12):

$$\frac{k''}{2^{N-1}} - \frac{1}{2^N} \leq \frac{t_k^{''}}{\frac{2}{L}T_C} < \frac{k''}{2^{N-1}} + \frac{1}{2^N}. \tag{12}$$

2.3. Exposure Time Ratio and Communication Speed

The proposed method indicates that although the intensity values obtained through Equation (1) and their intensity differences obtained through Equation (3) are different at different exposure time settings, the time delays calculated by Equation (4) can be the same. Thus, the proposed method allows setting a longer exposure time. It requires $4T_C/L$ to transmit one symbol, which is decoded by four consecutive line sensors. However, to avoid interference between symbols, their transmission duration must be longer than $4T_C/L$. In the following discussion, therefore, it is assumed that the symbol transmission lasts $(4+a)T_C/L$ $(a = 1, 2, \cdots)$ and received symbols are decoded by using four among $(4+a)$ line sensors. When η is set to $\frac{2}{pL}$ (Equation (7)) and an N-bit symbol is decoded using four among $(a+4)$ line sensors without errors, the communication speed is $N/(a+4)$ bits per line sensor. If an LED illumination is simultaneously exposed to all line sensors of the camera, the theoretical maximum speed becomes $NL/((4+a)T_C)$ bps.

Two cases are investigated when η is larger than $\frac{2}{pL}$.

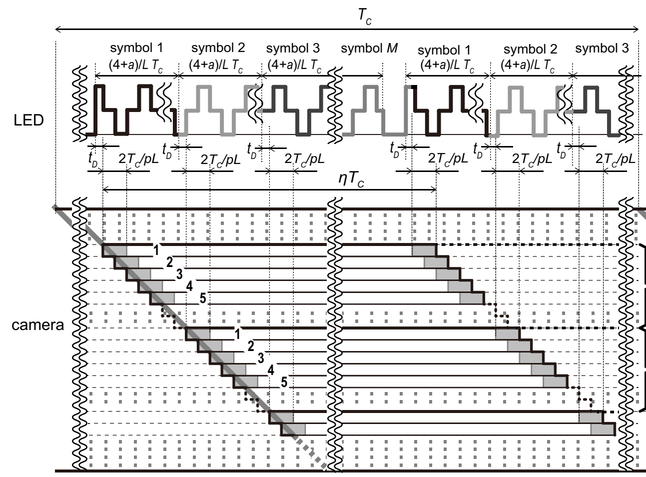

Fig. 4. Symbol transmission and reception for $(\eta - \frac{2}{pL}) \geq (4+a)/L$.

(a) If $(\eta - \frac{2}{pL}) \geq (4+a)/L$**.** As shown in Figure 4, if each line sensor receives the same signal in the first and last $\frac{2}{pL}$ periods of its exposure time, Equation (3) gives double the value of that when $\eta = \frac{2}{pL}$. Thus, Equation (4) gives the same time delay as that of $\eta = \frac{2}{pL}$. The number of symbols M to be transmitted in the $(\eta - \frac{2}{pL})T_C$ period is $\lfloor(\eta - \frac{2}{pL})L/(a+4)\rfloor$. As the symbols must be transmitted twice, $\lceil(\eta - \frac{2}{pL})L\rceil + (a+4)M$ line sensors are required to receive M symbols. Theoretically, $NM/(\lceil(\eta - \frac{2}{pL})L\rceil + (a+4)M)$ bits per line sensor can be achieved at maximum. Note that $\lceil x \rceil$ is the minimum integer larger than x.

(b) If $(4+a)/L > (\eta - \frac{2}{pL}) > 0$**.** A line sensor cannot receive the same symbol in the first and last $\frac{2}{pL}$ periods of its exposure time.

Thus, the requirement for the exposure time ratios is given as follows:

$$\eta = \frac{2}{pL}, 1 > \eta \geq \frac{4+a}{L} + \frac{2}{pL}. \tag{13}$$

Parameter	Experiment 1	Experiment 2	Experiment 3
Exposure time ratio η	1.91×10^{-3}	1.91×10^{-3}	3.24×10^{-2}, 1.01×10^{-1}
Measurement environment I_B	"dark", "fluo"	"dark"	"dark", "fluo"
Distance d (m)	0.03	0.03, 0.5, 1.0, 1.5	0.03
Number of pixels P_N	100, 300, 600, 1324	40	40
Bits per symbol N	6, 7, 8, 9, 10	6, 7, 8, 9, 10	3, 4, 5, 6, 7
Number of measurements	$1000 \times 5 \times 4 \times 2 = 40,000$	$2000 \times 5 \times 4 = 40,000$	$2000 \times 5 \times 2 \times 2 = 40,000$

Table 1. Experimental parameters and their values ("dark": no illumination, "fluo": fluorescent illumination).

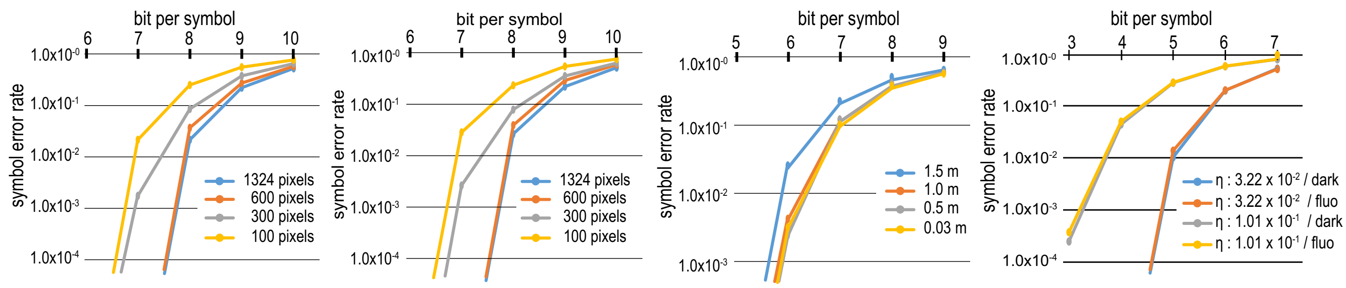

Fig. 5. Symbol error rates: (a) different numbers of pixels per line sensor in a dark environment; (b) different numbers of pixels per line sensor under fluorescent light; (c) different distances; and (d) different exposure time ratios.

3. EXPERIMENT

3.1. Overview

An LED (OptoSupply OSB56A5111A) connected to a function generator (NF Corporation WF1948) emitted an optimally modulated signal ($\eta = \frac{2}{L} = 1.91 \times 10^{-3}$ as shown in Figure 3(a)) for a 60 fps camera (Point Grey Flea3, USB 3.0, 1324×1048 pixels). A round diffuser was placed between them to expose as many line sensors as possible simultaneously.

A transmitted signal lasting one frame time (T_C) consisted of three parts: a nonsignal part longer than $6T_C/L$ to detect the start of the transmission at the receiver, a time-difference estimation part of duration $6T_C/L$, and an encoded symbol part.

Three experiments were conducted by changing measurement environments (I_B), number of pixels (P_N) per line sensor, exposure time ratios (η), distances (d) between LED and camera, as shown in Table 1. The proposed method decoded one symbol using four of six line sensors ($a = 2$ in Section 2.3). An LED was placed so that its image was captured at the center of the camera. The number of line sensors used for decoding was 50 (Experiment 2) or 1048 (otherwise).

3.2. Experimental Results and Discussion

The results of Experiments 1 and 2 confirm that the proposed method could achieve five-bit symbol transmissions without errors at the $d = 1.5$ m setting (Figure 5 (c)) and seven-bit transmissions without errors at the $d = 0.03$ m and $P_N = 600/1324$ setting (Figure 5 (a) and (b)). These corresponded to 0.833 and 1.17 bits per line sensor, and 54.0 kbps and 75.6 kbps using a 60 fps camera with 1080 line sensors, respectively.

The results of Experiment 3 (Figure 5 (d)) indicate that a larger η may improve the signal-to-noise ratio while decreasing the number of bits to be transmitted as symbols. For example, three- and four-bit symbols were transmitted without errors, which corresponded to 0.313 and 0.249 bits per line sensor at the $\eta = 3.24 \times 10^{-2}$/dark and $\eta = 1.01 \times 10^{-1}$/dark settings, respectively. Possible reasons for the deterioration are as follows.

1. A longer exposure time increased intensity values and made the differences obtained in Equation (3) relatively smaller.

2. The intensity of each pixel was represented as an eight-bit integer and smaller intensity differences might be affected by quantization errors.

4. CONCLUSIONS

An OCC technique for rolling-shutter cameras was proposed by extending a time-difference estimation method for global-shutter cameras. The method was confirmed to achieve better high-speed communication performance than existing OCC systems. The evaluations showed that although some performance deterioration was observed, the proposed method allowed longer exposure times, which could be useful for augmented reality. A future task is to implement the proposed technique on smartphones.

5. REFERENCES

[1] IEEE 802.15 WPAN^{TM} 15.7 Amendment - Optical Camera Communications Study Group (SG 7a), ," `http://www.ieee802.org/15/pub/SG7a.html` (retrieved Oct 7, 2017).

[2] T. Komine and M. Nakagawa, "Fundamental Analysis for Visible Light Communication System using LED Lights," *IEEE Transactions on Consumer Electronics*, pp. 100–107, 2004.

[3] R. Boubezari, H. L. Minh, Z. Ghassemlooy, and A. Bouridane, "Smartphone Camera Based Visible Light Communication," *Journal of Lightwave Technology*, vol. 34, no. 17, pp. 4120–4126, 2016.

[4] B. Zhang, K. Ren, G. Xing, X. Fu, and C. Wang, "SB-VLC: Secure Barcode-Based Visible Light Communication for Smartphones," *IEEE Transactions on Mobile Computing*, vol. 15, no. 2, pp. 432–446, 2016.

[5] W. Hu, H. Gu, and Q. Pu, "LightSync: Unsynchronized Visual Communication over Screen-Camera Links," in *ACM MobiCom'13*, Miami, FL, 2013, pp. 15–26.

[6] Y. S. Kuo, P. Pannuto, K. J. Hsiao, and P. Dutta, "Luxapose: Indoor Positioning with Mobile Phones and Visible Light," in *ACM MobiCom'14*, Maui, HI, 2014, pp. 447–458.

[7] M. Yasir, S. W. Ho, and B. N. Vellambi, "Indoor Positioning System Using Visible Light and Accelerometer," *Journal of Lightwave Technology*, pp. 3306–3315, 2014.

[8] N. Rajagopal, P. Lazik, and A. Rowe, "Visual Light Landmarks for Mobile Devices," in *IPSN '14*, Berlin, Germany, 2014, pp. 249–260.

[9] Z. Yang, Z. Wang, J. Zhang, C. Huang, and Q. Zhang, "Wearables Can Afford: Light-weight Indoor Positioning with Visible Light," in *ACM MobiSys'15*, Florence, Italy, 2015, pp. 317–330.

[10] C. Zhang and X. Zhang, "LiTell: Robust Indoor Localization Using Unmodified Light Fixture," in *ACM MobiCom'16*, New York, NY, 2016, pp. 230–242.

[11] S. Zhu and X. Zhang, "Enabling High-Precision Visible Light Localization in Today's Buildings," in *ACM MobiSys 2017*, Niagara Falls, NY, 2017, pp. 96–108.

[12] M. Sugimoto, H. Kumaki, T. Akiyama, and H. Hashizume, "Optimally Modulated Illumination for Rapid and Accurate Time Synchronization," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 505–516, January 2017.

[13] IEEE Standards Association, "IEEE Recommended Practices for Modulating Current in High-Brightness LEDs for Mitigating Health Risks to Viewers," `https://standards.ieee.org/findstds/standard/1789-2015.html` (retrieved Mar 26, 2016).

[14] H. Aoyama and M. Oshima, "Visible Light Communication Using a Conventional Image Sensor," in *IEEE CCNC 2015*, Las Vegas, NV, 2015, DOI:10.1109/CCNC.2015.7157954.

[15] C. Danakis, M. Afgani, G. Povey, I. Underwood, and H. Haas, "Using a CMOS Camera Sensor for Visible Light Communication," in *IEEE GLOBECOM 2012*, Anaheim, CA, 2012, pp. 1244–1248.