

Color-filter-free spatial visible light communication using RGB-LED and mobile-phone camera

Shih-Hao Chen and Chi-Wai Chow*

Department of Photonics and Institute of Electro-Optical Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan

*cwchow@faculty.nctu.edu.tw

Abstract: A novel color-filter-free visible-light communication (VLC) system using red-green-blue (RGB) light emitting diode (LED) and mobile-phone camera is proposed and demonstrated for the first time. A feature matching method, which is based on the scale-invariant feature transform (SIFT) algorithm for the received grayscale image is used instead of the chromatic information decoding method. The proposed method is simple and saves the computation complexity. The signal processing is based on the grayscale image computation; hence neither color-filter nor chromatic channel information is required. A proof-of-concept experiment is performed and high performance channel recognition is achieved.

©2014 Optical Society of America

OCIS codes: (230.3670) Light-emitting diodes; (060.4510) Optical communications; (060.4080) Modulation.

References and links

1. C. W. Chow, C. H. Yeh, Y. Liu, and Y. F. Liu, "Digital signal processing for light emitting diode based visible light communication," *IEEE Photon. Soc. Newslett.* **26**, 9–13 (2012).
2. H. H. Lu, Y. P. Lin, P. Y. Wu, C. Y. Chen, M. C. Chen, and T. W. Jhang, "A multiple-input-multiple-output visible light communication system based on VCSELs and spatial light modulators," *Opt. Express* **22**(3), 3468–3474 (2014).
3. W. Y. Lin, C. Y. Chen, H. H. Lu, C. H. Chang, Y. P. Lin, H. C. Lin, and H. W. Wu, "10m/500 Mbps WDM visible light communication systems," *Opt. Express* **20**(9), 9919–9924 (2012).
4. Z. Wang, C. Yu, W. D. Zhong, J. Chen, and W. Chen, "Performance of a novel LED lamp arrangement to reduce SNR fluctuation for multi-user visible light communication systems," *Opt. Express* **20**(4), 4564–4573 (2012).
5. Z. Wang, W. D. Zhong, C. Yu, J. Chen, C. P. Francois, and W. Chen, "Performance of dimming control scheme in visible light communication system," *Opt. Express* **20**(17), 18861–18868 (2012).
6. C. H. Yeh, Y. L. Liu, and C. W. Chow, "Real-time white-light phosphor-LED visible light communication (VLC) with compact size," *Opt. Express* **21**(22), 26192–26197 (2013).
7. J. Y. Sung, C. W. Chow, and C. H. Yeh, "Dimming-discrete-multi-tone (DMT) for simultaneous color control and high speed visible light communication," *Opt. Express* **22**(7), 7538–7543 (2014).
8. C. Danakis, M. Afgani, G. Povey, I. Underwood, and H. Haas, "Using a CMOS camera sensor for visible light communication," *Proc. IEEE Workshop on Opt. Wireless Comm. 2012 (OWC'12)*, 1244–1248.
9. I. Takai, S. Ito, K. Yasutomi, K. Kagawa, M. Andoh, and S. Kawahito, "LED and CMOS image sensor based optical wireless communication system for automotive applications," *IEEE Photon. J.* **5**(5), 6801418 (2013).
10. S. Sugawa, N. Akahane, S. Adachi, K. Mori, T. Ishiuchi, and K. Mizobuchi, "A 100 dB dynamic range CMOS image sensor using a lateral overflow integration capacitor," *Proc. ISSCC 2005, Papers 19*.
11. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.* **60**(2), 91–110 (2004).
12. P. Ge, P. Yin, H. Wang, and T. Chang, "Image matching technology in high power LED's eutectic welding," *Opt. Express* **22**(11), 13531–13540 (2014).

1. Introduction

Wireless communication plays an important role in our daily lives. Light emitting diode (LED) based visible light communication (VLC) can be a promising choice [1–3]. LED is suitable for VLC since it can be modulated at a relatively high speed when compared with traditional light sources [4–7]. Like LED, image sensor or camera has been widely used in our daily lives. The majority of new generation mobile-phones have embedded

Complementary Metal-Oxide-Semiconductor (CMOS) cameras allowing users to capture photos and videos. Hence, VLC using LED lamps and the already existing mobile-phone and vehicle cameras is attractive and provides a low-cost wireless communication. VLC using CMOS sensor and vehicle camera have been reported in [8, 9] respectively. Ref [8]. only focuses on operating the CMOS sensor at a much higher data reception rate than the camera frame rate using the rolling shutter effect of the sensor. Ref [9]. fabricates an image sensor with specific pixels for imaging and high speed communication. However, this sensor is not commercially available and may increase the system complexity and costs. As red-green-blue (RGB) LED signboards are widely used, it is believed that this $n \times n$ RGB LED array can act as the VLC transmitter (Tx). The receiver (Rx) can be the embedded CMOS camera in mobile-phone. The two dimensional (2D) spatial Tx and Rx can improve the transmission capacity by transmitting different data at different LEDs. This can be considered as a multiple-input and multiple-output (MIMO) VLC transmission. In addition, as the LED array is consisted of RGB LEDs, wavelength division multiplexing (WDM) can be used to further increase the transmission capacity. In this $n \times n$ WDM LED-based VLC system, the key step of decoding the received RGB signals is to distinguish the three color channels. Traditionally, Rx should equip with three color filters, and three times more memories are required to store the RGB color information than a grayscale image of the same dimension [8]. Moreover, most of the decoding algorithm is based on using color-filters to discriminate the three color channels. The chromatic information can be easily affected or lost owing to the image sensor saturation caused by the high intensity optical signal [10]. This means the pixels under high intensity optical illumination are turned into a big “white spot” without knowing which color channel is used.

In this work, a novel color-filter-free VLC system using RGB-LED and mobile-phone camera is proposed and demonstrated for the first time. A feature matching method of the received grayscale image is used instead of the chromatic information decoding method. The feature matching method is based on the scale-invariant feature transform (SIFT) [11] algorithm. In this scheme, image features from the received grayscale image are extracted. They are then used to compare with the golden reference features which are extracted from the training sequence. The signal processing is based on the grayscale image computation; hence neither color-filter nor chromatic channel information is required to decode the received signals. Moreover, the proposed method is simple and saves the computation complexity; hence the matching based on knowing the geometry features of the whole image as reported in [12] is not required. Furthermore, the proposed feature matching can identify the specific color diffusion profile; hence it can solve chromatic information lost issue (pixels under high illumination are turned into a big “white spot” without knowing which color channel used) owing to the image sensor saturation by high intensity light. A proof-of-concept experiment is performed and high performance channel recognition is achieved using the proposed scheme.

2. Architecture and principle

The architecture of the proposed color-filter-free RGB VLC system is shown in Fig. 1. The Tx equips with a 2D RGB LED array. The input data is first scrambled and bit-mapped into the 2D RGB-LED array. In our proof-of-concept experiment, the RGB LED array size is 3×3 ; hence the total number of $3 \times 3 \times 3$ channels can be used in WDM operation. On-off keying (OOK) is used to demonstrate the proposed system. Proper power level is selected to control the LED output optical powers. Then training sequences are added in front of the data sequences. The training sequence is used to setup the golden reference feature for decoding. The training sequence is used to turn on each color channel sequentially. A color channel is at “ON” state for 1/30 s (if 10% pulse width modulation (PWM) is used, the LED is turned on for 1/300 s), and then other color channel is at “ON” state for 1/30 s. The sequence pattern is R, G, B, “all off”. This pattern will be repeated 3 times to synchronize the Tx and Rx. PWM is used for adjusting the dimming level of each LED. 10% duty cycle of PWM is for the LED “ON” state and 0% duty cycle is for the LED “OFF” state. The data rate of the Tx is 30 Hz,

which is the same as the camera frame rate. We use the same duty cycle of PWM for each color LED. Using longer duty cycle of PWM (less dimming) can increase the output optical power and enhance the VLC transmission distance. Finally in the Tx, the 2D RGB LED array transmits the spatial VLC signals to the Rx. During the training sequence process, the image feature of each color channel is extracted. The extracted image feature is used as golden reference feature of each color channel. After the transmission of training sequence, the payload data will be transmitted. The golden reference features extracted from the training sequence will be compared with the image features extracted from the payload data sequences.

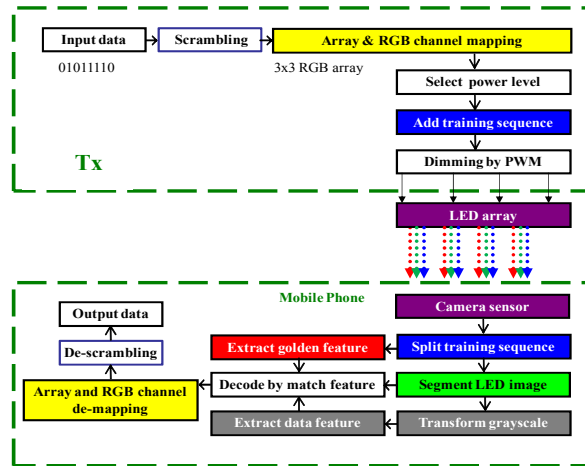


Fig. 1. Functional block diagram of the proposed RGB-VLC system.

iPhone 4s mobile-phone camera is used as the CMOS image sensor in the proof-of-concept demonstration. It records the 2D signals into video data, which contains 30 frames per second. The training sequence is separated and processed to extract the golden reference feature of each color channel. In the Rx, the received image is divided into nine image segments. Each segment contains only one RGB-LED image inside. The image features are extracted from the nine image segments instead of extracting the whole received image feature. This saves the cost for large scale computation of octave/scale feature extraction works. The image signal processing is based on the grayscale image computation. Hence, the proposed system is a color-filter-free system. After that, the extracted image features from the data sequences will be matched with the golden reference feature. The golden reference is extracted from the training sequences. If the number of matched image features is above the pre-defined threshold, data “1” is decoded. Otherwise, the data “0” is decoded. The decoded data is de-mapped, de-scrambled and finally the output data can be obtained.

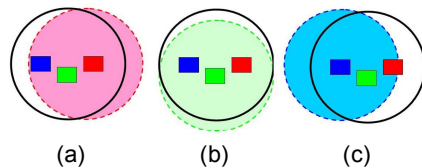


Fig. 2. Schematic top-view of the LED light spots. (a) Only red, (b) green, or (c) blue color chip is turned on in LED package.

Then, we discuss the principle of the image feature matching using SIFT. Each RGB-LED is composed of red, green and blue chips emitting three color channels. Every color channel can form a specific light spot within one single LED. Each specific light spot has individual light diffusion profile. Rx can extract the image feature of individual light diffusion profile by

using the SIFT algorithm. The individual light diffusion profile is due to the geometric packaging difference of each color chip within a LED. The extracted image feature of individual light diffusion profile is different. Hence, this can be used to identify the color chip without the need of the chromatic information. Schematic LED light spots shown in Fig. 2 are used for the explanation. The black circle conceptually illustrates the LED package size. Figure 2(a) shows only the red color chip is turned on and the red light spot emitted by the LED will shift to the right side. When the green or blue color chip is turned on, the light spot will shift down or to the left accordingly as shown in Fig. 2(b) and 2(c) respectively. Hence, the extracted image feature of individual light diffusion profile is different. That can be used to identify which color channel is transmitted data without the need of color-filters.

The SIFT algorithm is used to extract image feature of the received images. The image feature is extracted from the light distribution profile characters. The feature extracting algorithm is based on Difference of Gaussian (DoG) filter. Here, we introduce its operation principle.

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (1)$$

Equation (1) illustrates the 2D Gaussian filter. The received image $I(x,y)$ in Eq. (2) is convolved with the 2D Gaussian filter to get the Gaussian-blurred image.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2)$$

By the difference of the successive Gaussian-blurred images with different standard deviation (multiply by k) shown in Eq. (3), key points are taken as maxima/minima of the DoG that occurs in interested scales. In order to speed up the searching time, we choose k from 1 to 4.

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3)$$

Hence, the image feature can be investigated. As above mentioned, the Rx extracts the image feature from the training sequences as golden reference feature of each color channel. Then, the extracted image features from the data sequences will be matched with the golden reference feature. It matches with the three color features to find the most likely one in order to decide which color channel is turned on. If the number of matched image features is above the pre-defined threshold, the data “1” is decoded. Otherwise, the data “0” is decoded.

3. Experiment, results and discussion

A proof-of-concept experimental demonstration is performed. Figure 3(a) shows the photograph of the Tx, consisting of 3 x 3 RGB LED array and a self-design controller board. The controller board includes the function modules described in Fig. 1. We first extract the golden reference feature from the training sequences. Refer to the Table 1, the three color channels have different number of extracted image features. 48, 83 and 71 image features are extracted from the red, green and blue color channels respectively in average at a transmission distance of 50 cm. 39, 68 and 57 image features are extracted in average at a transmission distance of 100 cm as also shown in Table 2. The scale-invariant feature transform (SIFT) algorithm is not related to the transmission distance; hence increasing the transmission distance will not increasing the computation complexity. The main limitation of the proposed system is the number of pixels (detected at the image Rx) occupied by each LED. As shown in Table 2, the matched image features decrease when the transmission distance increases from 50 cm to 100 cm. This is because fewer detected pixels are represented as an LED. This degrades the feature searching performance. Figure 3(b) shows, for example, the red color channel of the LED is on at time 1, while Fig. 3(b) shows the same red color channel is on at time 2. The blue line connects (used for illustration) the matched image features between Figs. 3(b) and 3(c). There are 37 image features matched. On the other hand, Fig. 3(d) shows the red color channel of the LED is on at time 1, and Fig. 3(e) shows the green color channel

is on. The blue line connects the matched image features between Figs. 3(d) and 3(e), showing there are only 2 image features matched. Comparing the two results, it shows that the feature matching method can efficiently distinguish the different color channel in the grayscale domain. In the experiment, we have experienced the color saturation issue in Figs. 3(b) and 3(c), and the Rx cannot show the actual color of the LED. By applying the proposed feature matching scheme, it can identify the white spot shifting by comparing with the golden reference.

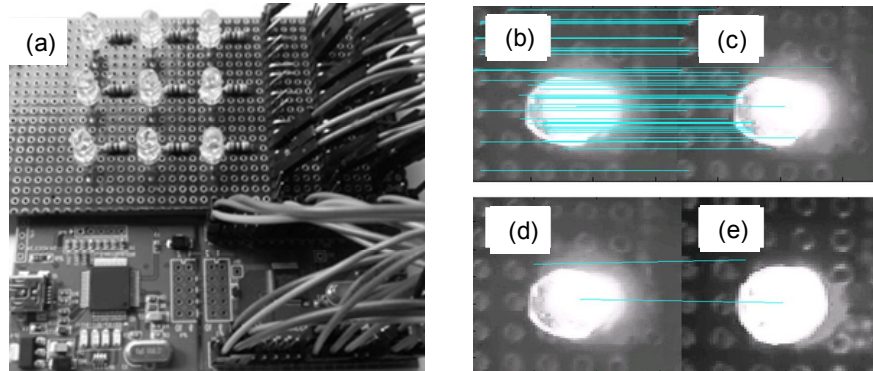


Fig. 3. (a) Photograph of the Tx, showing the 3 x 3 RGB LED array and the control board. Feature matching of LED images: (b) red color channel on, (c) red color channel on, (d) red color channel on, and (e) green color channel on. Blue lines illustrate the matched image features.

Table 1. Average Number of Extracted Features of 3 Color Channels @ Transmission Distance 50cm/100cm

LED emit Color	Red	Green	Blue
Average number of extracted features @ 50cm	48	83	71
Average number of extracted features @ 100cm	39	68	57

Table 2. Average Number of Matched Feature of 3 Color Channels @ 50cm/100cm

LED emit Color	Red	Green	Blue
Average matched features with Red reference @50cm	45	3	5
Average matched features with Green reference @50cm	7	79	5
Average matched features with Blue reference @50cm	6	4	59
Average matched features with Red reference @100cm	29	5	4
Average matched features with Green reference @100cm	8	41	7
Average matched features with Blue reference @100cm	9	6	36

Moreover, the average cross feature matching results are verified and shown in the Table 2. At 50 cm distance, when red color channel is on, 45 image features are matched with the red color golden reference feature in average. Only 3 image features are matched with the green color golden reference feature and 5 image features are matched with the blue color golden reference feature in average. The green and blue color channels have 79 and 59 matched features in average with their corresponding golden reference features. Moreover, the cross feature matching results at a 100 cm transmission distance is shown in the Table 2. These show that the feature matching method can efficiently distinguish the 3 different color channel in the grayscale domain even at 100 cm.

That result is clear for setting the pre-defined threshold to decode the data. We choose half of the average number of extracted features (in Table 1) to be the pre-defined threshold. For example, 48 image features is extracted from the red color channel. Then, we set 24 to be the pre-defined threshold of the red color channel. Table 3 shows the experimental results of the correct detection rate of each color channel. The red, green and blue color channels correct detection rates are 99.97%, 99.98% and 99.98% respectively. Four states (including the turn off mode) can be used to encode dual bits in the RGB-VLC proposed system.

As long as the light spot of the image can be extracted, other modulation formats can be used in the proposed scheme. The proposed VLC system based on 2D spatial $n \times n$ LED array and mobile-phone camera can be considered as a MIMO transmission. Besides, as RGB LEDs are used, WDM is used to further increase the transmission capacity. In this proof-of-concept experiment, we record the video data with 30 frames per second (in mobile-phone). Since $3 \times 3 \times 3$ LED array can transmit $3 \times 3 \times 3$ bits per frame, 810 bits per second is the throughput of the proposed system. As the proposed system describes signals in spatial domain. If the spatial resolutions of the Tx and Rx are increased to represent more signal channels, the system capacity can be further enhanced. Here, the proposed system is based on iPhone 4s mobile-phone camera (Rx), which supports 30 frames per second video recording. In the new generation of iPhone 6 Plus, 240 frames per second can be supported in slow motion video recording. The system capacity could be enhanced by 8 times. However, using high speed cameras will increase the system cost. The experiment is performed in the indoor environment with the ambient light color temperature of 5700 K. The experimental measured bit-error-rate (BER) is 1.9×10^{-4} in average at transmission distance of 50 cm. This satisfies the forward error correction (FEC) requirement.

Table 3. Average Detection Rate of Three Different Color Channels @ 50cm/100cm

LED emit Color	Red	Green	Blue
Average matched features with Red reference @ 50cm	99.97%	0.01%	0.01%
Average matched features with Green reference @ 50cm	0.01%	99.98%	0.01%
Average matched features with Blue reference @ 50cm	0.02%	0.01%	99.98%
Average matched features with Red reference @ 100cm	99.96%	0.02%	0.01%
Average matched features with Green reference @ 100cm	0.01%	99.97%	0.04%
Average matched features with Blue reference @ 100cm	0.03%	0.01%	99.95%

4. Conclusion

A novel color-filter-free VLC system using RGB-LED and mobile-phone camera was proposed and demonstrated. By using the color channel feature matching method based on the SIFT algorithm, the system experimentally achieved correct detection rate above 99.97% in average. Furthermore, the proposed feature matching can identify the specific color diffusion profile; hence it can solve chromatic information lost issue owing to the image sensor saturation. A proof-of-concept experiment was performed (throughput was 810 bits per second) and high performance channel recognition was achieved using the proposed scheme. The measured BER was 1.9×10^{-4} , and satisfied the FEC requirement.

Acknowledgments

This work was supported by Ministry of Science and Technology, Taiwan, MOST-103-2221-E-009-030-MY3, MOST-101-2628-E-009-007-MY3, Aim for the Top University Plan, Taiwan, and Ministry of Education, Taiwan.