附件1-5 论文5 本项目成员关伟鹏为论文作者

Signal Processing: Image Communication 75 (2019) 128–140

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

The optical fringe code modulation and recognition algorithm based on visible light communication using convolutional neural network☆

Heng Zhang [a], Yongjun Li [a], Weipeng Guan [b,*], Jingyi Li [b], JieHeng Zheng [a], Xinjie Zhang [c]

[a] School of Computer Science and Technology, South China University of Technology, Guangzhou, Guangdong 510640, China
[b] School of Automation Science and Engineering, South China University of Technology, Guangzhou, Guangdong 510640, China
[c] School of Electronic and Information Engineering, South China University of Technology, Guangzhou, Guangdong 510640, China

ARTICLE INFO

Keywords:
Visual communication
Image communication
RGB-LED
Optical fringe code (LED-OFC)
Complementary metal–oxide–semiconductor (CMOS) image sensor
Convolutional neural network (CNN)
Modulation and recognition

ABSTRACT

Recently, visible light communication (VLC) based on complementary metal–oxide–semiconductor (CMOS) sensor has been widely studied, and most of the research uses the modulation and demodulation method, which modulates the light emitting diode (LED) light to transmit data and demodulates bright and dark stripes on the image captured by the CMOS sensor to get data. However, the method have some defects. Firstly, as the distance increases to a certain extent, the data frame structure will be partially lost. Secondly, the image captured by the CMOS sensor must be strictly synchronized, which is hard to guarantee. Thirdly, the focus of recent related research is mainly on the real-time nature of communication, which is difficult to achieve due to the complex image processing methods at this stage. What is more, for many application scenarios recognizing the LED information in the image captured by the CMOS sensor is enough. So, in this paper, we introduce the RGB-LED and propose an optical fringe code (LED-OFC) modulation and recognition algorithm based on VLC using convolutional neural network (CNN). The RGB-LED is modulated to assign different features to the LED-OFC instead of transmitting data. Then the CNN is employed to recognize LED-OFC with different features. The experiment results show that both the recognition accuracy, the recognition amount, the maximum recognition distance and the robustness are greatly improved by the proposed method compared with the traditional modulation and demodulation method, which has broad application prospects.

1. Introduction

Due to the advantages of low energy consumption and long operational life, light emitting diode (LED) has become the preferred illumination source and is widely used in various occasions such as indoor lighting, advertising signs, traffic lights and street lights in recent years. In addition, LED light can be modulated at frequencies up to hundreds of MHz which is undetectable by the human eye [1,2], that makes visible light communication (VLC) based on LED possible. And lots of research has been conducted in VLC field [3–7]. In [3], the author researched the key challenges and technologies for applying VLC in a 5G network especially in downlink communications, and proposed a heterogeneous multi-layer 5G cellular architecture with a control plane and user plane separation scheme which is able to guarantee both low cost and high throughput. The cost of LED infrastructure is expected to further decrease by 55 percent [3], and the average throughput can maintain about 70 Mbps when the number of users increases from 1 to 6, while the throughput of wireless fidelity (WiFi) decreases from 30 Mbps to 2 Mbps when the number of users increases from 1 to 6. The proposed scheme proves that VLC has enormous potential to become a main feature in the 5G area. In [4], the key technologies and shortcomings of indoor VLC systems are thoroughly studied. A mobile angular diversity optical receiver detector model with both specular and diffuse reflections is proposed which increase the likelihood of line of sight (LOS) among transmitters and receivers. It is also shown that in rooms with reflection coefficients of 0.5 or above, diffuse reflection scenarios yield a 50% gain or in system throughput compared to LOS case. In [5], a light fidelity (LiFi) transmission system with a transmission distance of 10m and data rate of 25 Gbps based on a two-stage injection-locked 680 nm vertical-cavity surface-emitting laser (VCSEL) transmitter is proposed, which achieves bit error rate (BER) of 10^{-6} and clear eye diagram even when the distance is 10m. The experiment results prove that the proposed LiFi transmission system has great potential as a replacement for wireless networks. So, transmitting data by visible light is not surprising. Compared with

☆ No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.image.2019.04.002.

* Corresponding author.
E-mail address: augwpscut@mail.scut.edu.cn (W. Guan).

https://doi.org/10.1016/j.image.2019.04.002
Received 5 September 2018; Received in revised form 2 April 2019; Accepted 2 April 2019
Available online 5 April 2019

traditional radio frequency (RF) wireless communications, the advantages of VLC are listed as follows. Firstly, the visible light portion of the electromagnetic spectrum is unlicensed and unregulated, and LED is more energy efficient than other radio wave transmitters [8], which is a significant cost saving. Secondly, the line of sight (LOS) property of VLC provides an inherent transmission security [9], which is important in communication. Thirdly, via modulating the LED, VLC can achieve both illumination and data transmission, which does not require any additional transmitted power [10]. So, considering VLC as a complementary to traditional RF wireless communications is feasible and attractive.

In a typical VLC system, LED is used as transmitter. As mentioned before, LED can switch between on and off at a very fast rate, which is enough to transmit 1/0 data bit without visible flicker. There are two choices can be employed as receiver: photodiodes (PDs) [11], and image sensors [12]. Due to the features of low cost and high transmission data rate, the PD is widely used as receiver, but PD is not conducive for practical applications and direct line of sight during signal detection is needed [13], the background light and reflected light are non-negligible factors which make the system unstable [14], and therefore the maximum transmission distance is short. In contrast, although the data transmission rate is slower than PD, using image sensor as the receiver is more stable since the ambient light has less inference [15]. The direct light projected on the image sensor is considered as the foreground while the ambient light is considered as background, and it is easy to separate them by image processing methods or others, hence the maximum transmission distance is longer. In addition, a multitude of smart devices has been equipped with image sensors, including smartphones, laptops and automobiles. So, from the perspective of system stability and ease of use, VLC based on image sensor is a better choice.

Using traditional modulation and demodulation method is the most intuitive way to transmit data in VLC. The LED is modulated by some modulation method to carry 0/1 data, and the transmitted data is captured by image sensor in the form of bright and dark stripes (bright stripes represent 1 and dark stripes represent 0) on the image. Then a series of image preprocessing operations needs to be performed on the captured image to get the individual stripes. Finally, demodulation in a corresponding manner is used to obtain the transmitted data. Many studies on VLC based on image sensor using modulation and demodulation method have been carried out. In [16], B. Zhang et al. proposed a secure system for barcode-based VLC system between smartphones, and the results show that the system achieves high level security and NFC-comparable throughput. The author tested their scheme for different versions of quick response (QR) code and achieved throughput of 70 kbps without the security transmission scheme, and the throughput dropped to 10 kbps when using the security transmission scheme, while the throughput of a typical NFC scheme is from 106 kbps to 424 kbps [17]. Moreover, the throughput increased to 200 kbps when replace QR codes with color barcodes. However, the sender of the system uses an LED in the smartphone which would cause uneven light output, further, the error rate would increase due to the "blooming" effect [15] of image sensor caused by the uneven light output. Since the uneven light output will result in excessive illumination in the local area of the image sensor, then the charge will overflow from saturated pixel to adjacent pixel, finally a bright light spot will appear on the image captured by the image sensor, which makes demodulation impossible. And the maximum transmission distance is short (about 10 cm to 20 cm). In [18], the author uses an LED panel light to reduce the possibility of overexposure and avoid the blooming effect of the camera. The method achieves error-free performance in symbol error rate (SER) test. But the maximum transmission distance is too short (about 4 cm), when the distance is larger the data packet might be partly lost. And the transmission data rate is only 4 Kbits/s, which is low. In [19], C. W. Chen et al. packs the data into a packet to transmit and achieves 64 bits payload at one time. However, the maximum transmission distance is only a few dozen centimeters, the transmission data which is rendered as an image captured by the image sensor can only be demodulated correctly once the complete data packet is received, but it is hard to guarantee. In [20,21], the author uses a light-panel with size of 577 mm x 878 mm as transmitter and mobile-phones as receiver, and the experiments compare the success rate (or bit-error-ratio) under different NRV (noise-ratio value), transmission rate and transmission distances. The experiment results show that both mobile-phones under test can achieve success rate > 96% even when the transmission distance is up to 200 cm when NRVs are low, and even at high NRV (70.21%) the FEC threshold (bit-error-ratio = 3.8×10^{-3}) can be achieved at the data rate of 480 bit/s and the transmission distance of 150 cm. However, this does not solve the problem, since the size of the LED is much bigger than it in the above work, and it is unfair to compare their maximum transmission distance directly. Moreover, synchronization is required for all the above work, since premature data reception may result in data redundancy, while late data reception may result in incomplete data [15].

In our previous work [15], in-depth research has been conducted on VLC based on image sensor using traditional modulation and demodulation method. We find that traditional modulation and demodulation method still have the following problems. Firstly, the modulation and demodulation method is severely limited by the distance between transmitter and receiver which can only reach 10 cm to 50 cm, since the transmission data frame structure would be partially lost and the stripes of the resulting image in the image sensor will be blurred [15] which increases the difficulty of preprocessing as the distance increases (The experiment later in this paper will prove this). Secondly, for the modulation and demodulation method, when receiving data, the image captured by the image sensor must be strictly synchronized. Since premature data reception may result in data redundancy, while late data reception may result in incomplete data [15]. However, synchronization is difficult to guarantee, which will have an impact on robustness and accuracy. Thirdly, complex image processing methods including grayscale conversion and image segmentation, column pixel selection, smoothing and extinction ratio enhancement, and threshold detection [15] before demodulation reduce the transmission rate, which makes real-time communication difficult to achieve. What is more, for many application scenarios of VLC based on image sensor, such as positioning LED-ID recognition, identification of LED traffic lights, LED optical stripe code recognition and optical stripe code recognition and access, high real-time performance is not required, recognizing the LED image captured by the image sensor is enough, then the related information can be obtained once the image is recognized correctly. Therefore, ensuring the accuracy and robustness of recognition is the most important for most application scenarios of VLC based on the image sensor. However, VLC based on image sensor using traditional modulation and demodulation method is not suitable for these scenarios due to the above disadvantages.

Different from traditional modulation and demodulation method, in VLC based on image sensor there is another way to transmit data called modulation and recognition method [14], which convert the problem to a classification and recognition problem. The role of modulation step here is completely different from its role in traditional modulation and demodulation method. In modulation and recognition method, the purpose of modulation is to make the bright and dark stripes on the image captured by image sensor have as many different features as possible instead of transmitting the actual data, and different images can be given an independent identifier. Then the captured image with different stripe features can be recognized separately. Once the captured image with different stripe features is recognized correctly, we can get the related information by querying the database using the independent identifier.

And in our previous work [14], the modulation and recognition method is adopted in VLC based on the image sensor. Each LED is given a unique identification (LED-ID), which is associated with related

location information, then the demodulation step can be replaced by recognition step. By image preprocessing methods, some features are selected manually. Then machine learning methods, to be specific, Fisher Classifier and Linear Support Vector Machine are used to get the correct LED-ID by different features. The results show that the method can classify 1035 different LED-ID with high accuracy which is over 95%, and this is enough for large-scale positioning. However, there are still some limits. Firstly, to improve the recognition accuracy, the features need to be selected and extracted manually, and various image processing methods are required, which increases the complexity of the system and the actual recognition effect is not guaranteed. Secondly, Pulse-Width Modulation (PWM) method is used to change the duty-ratio in a large range, which would affect light intensity and decrease the recognition accuracy indirectly, especially when the distance is very large. Thirdly, the minimum frequency interval is different between the high frequency range and the low frequency range, but it is not considered, which decreases actual recognition accuracy. Finally, limited by the classification ability, the number of LED-ID is far from enough, which is a limitation for a wider range of applications.

So, in this paper, RGB-LED is introduced and modulated by different parameters which is to assign more different features to the optical fringe code (LED-OFC) captured by the complementary metal–oxide–semiconductor (CMOS) sensor rather than transmitting data. And an optical fringe code (LED-OFC) modulation and recognition algorithm based on VLC using convolutional neural network (CNN) is proposed, moreover, a VLC-CNN system is built to verify the feasibility. Specifically, the phase difference is first introduced to control the phase of different LEDs, which will produce mixed lights of different colors and make the stripes of the LED-OFC have different colors and color change cycle features. Except the phase difference, frequency and distance parameters are also used to modulate the RGB-LED which will make the width and the number of the stripe of the final LED-OFC different. In summary, the stripes of the final LED-OFC captured by the CMOS sensor will have different width, number, color and color change cycle features. Then a CNN architecture is introduced after LED-OFC is captured, which is employed to recognize different LED-OFCs with high accuracy and robustness. Furthermore, an LED-OFC-INFO database is built to get the related information when the LED-OFC is recognized correctly. Compared with traditional modulation and demodulation method and our previous work, the advantages of the proposed method are: Firstly, the proposed method is end-to-end style and no additional pre-processing methods are required, while traditional modulation and demodulation method requires a lot of preprocessing, equalization, and noise reduction operations, and the final demodulation result is easily interfered by various factors. Secondly, the features of the LED-OFC are selected automatically by the CNN, while that need to be selected and extracted manually in our previous work, which has an impact on the robustness and accuracy of the system. Thirdly, the proposed method can greatly improve the maximum recognition distance and the recognition amount with high accuracy and robustness compared with all recent related work, while synchronization is not required. And the experiment results also prove these advantages, the maximum recognition distance reaches 3.5 m which is dozens of times than that in recent related works, and the recognition number of LED-OFCs is more than 50 thousand, which is also far more than our previous work. And our contributions are as follows:

(1) RGB-LED is introduced which will produce three kinds of light separately. By assigning different parameters to different LEDs, the final LED-OFC will have more different features.
(2) Phase difference is first introduced which will make the final LED-OFC have different color and color change cycle features, and that further increased feature quantity of the LED - OFC.
(3) To the best of our knowledge, it is the first time that CNN is introduced into VLC field, which further improves the recognition distance, accuracy and robustness.
(4) To verify the feasibility of the proposed scheme, we explore and analysis the effects of various parameters on the final recognition performance, and this increases the system availability.

The remainder of this paper is organized as follows: System principle is described in Section 2. Experiment exploration and analysis is in Section 3. Finally, in Section 4, the conclusion is provided.

2. System principle

2.1. Image sensor detection

2.1.1. Rolling shutter mechanism

There are two types of image sensors: Charge Coupled Devices (CCD) sensor and CMOS sensor. The principle of CCD sensor is called “Global Shutter Mechanism”, which means that all pixels in the CCD sensor are exposed at the same time, and at the end of exposure, all pixels are read out at the same time too, as shown in Fig. 1(a). Different from the CCD sensor, for a CMOS sensor, the exposure and read operations are performed row by row, and the delay time of adjacent rows is equal to the read time, which is called “Rolling Shutter Mechanism” [15,22,23], as shown in Fig. 1(b). The period of one frame is defined as the difference between the time when the first row is reset and when the last row is read out. During this period, by turning the LED on and off with a high frequency, bright and dark stripes will appear on the image, which represent the modulated data, and several bits can be transmitted in one frame image. Although the data rate of the rolling shutter mechanism is as low as several Kbit/s [24], as mentioned before, it is enough for many application scenarios such as positioning LED-ID recognition, identification of LED traffic lights, LED optical stripe code recognition and optical stripe code recognition and access.

2.1.2. Camera Requirement

When using the CMOS sensor as receiver, two important properties must be considered: exposure time, and camera sensitivity (ISO). During exposure time, pixels accumulate charge until saturation when illuminated by light. Since the LED is also the role of illumination, the charge of a pixel will reach saturation in a short period of time, then the charge would spill over into the nearby pixel and cause the nearby pixel to accumulate charge. So, the exposure time should be controlled otherwise the width of the stripes would decrease due to long exposure time. As for ISO, it indicates the number of photons to make the pixels reach saturation. The higher the ISO, the fewer photons are needed. Therefore, as the ISO increases, the probability that pixels reach saturation increases, then the width of stripes also increases.

2.2. LED-OFC modulation

Similar to the related works, an on–off keying (OOK) scheme [25] is employed to modulate the LED, as shown in Fig. 2, the data bits 1/0 are encoded and transmitted via turning the LED on/off periodically using OOK scheme. However, different with the related works, to produce more LED-OFCs with different features, RGB-LED is introduced in this paper, and a lot of parameters are considered in the modulation step, which will be described in detail in the following subsections.

2.2.1. LED-OFC color control

In related works [14,16,18,19], the light produced by the LEDs is white, and the principle is as follows: In the process of perceiving objects around us, our eyes go through a process called the sum of time, which is also called “persistence of vision”. In the process, the eye accumulates photons until it is saturated. This period is called t_c:

$$\zeta = I * t \left(t \le t_c \right). \tag{1}$$

As shown in Eq. (1), I is the intensity of the stimulus (LED project's color stimulus). t_c is the critical duration. With the increase of it,

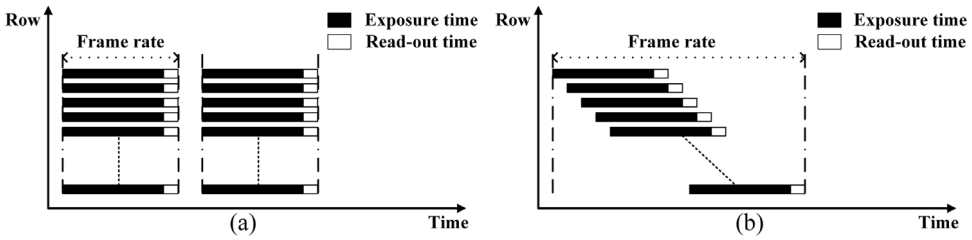

Fig. 1. The principle of two types of image sensors. (a) Global Shutter Mechanism of the CCD sensor; (b) Rolling Shutter Mechanism of the CMOS sensor.

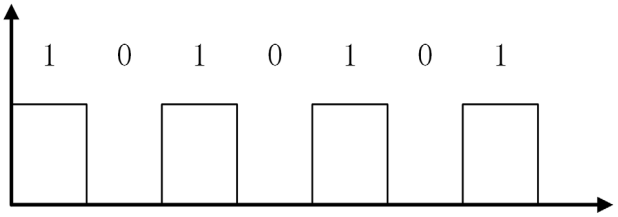

Fig. 2. The OOK scheme.

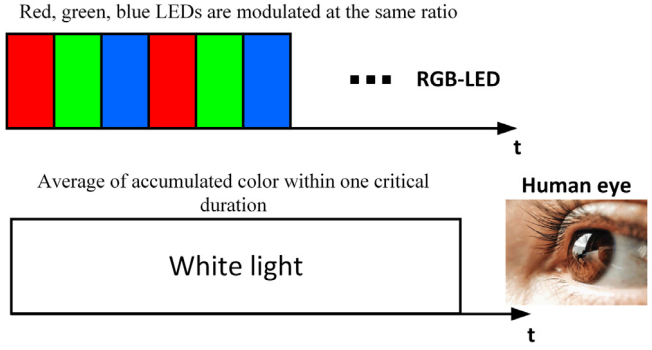

Fig. 3. The principle of white light in the human eye.

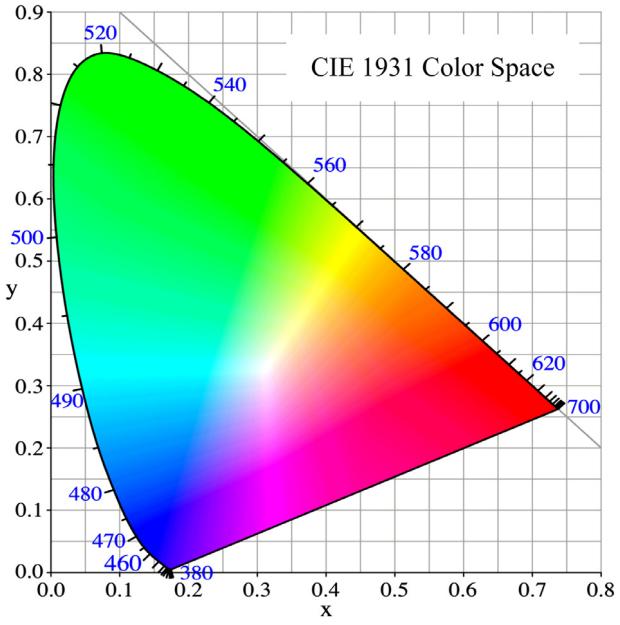

Fig. 4. The CIE 1931 RGB color space.

once the threshold is reached, the additional stimulus light will not be perceived by the visual system. The perception of color by the human eye is the average of the stimulus within the critical duration. According to Bloch's law, the color ψ is:

$$\psi = \frac{\int I_R(t)dt + \int I_G(t)dt + \int I_B(t)dt}{t}. \quad (2)$$

In Eq. (2), $I_R(t)$, $I_G(t)$, $I_B(t)$ are the intensity functions of red, green and blue light respectively. For RGB-LED, if the three types of light are emitted at the same ratio, the human eye will perceive the white light in the critical duration for the sum of the time. And the schematic is shown in Fig. 3.

However, if the final light of the LED is white, the final image produced by the CMOS sensor can only have white and black stripes, which does not make good use of the color features of the LED-OFC. By mixing different ratios of red, blue and green light according to CIE 1931 RGB color space [26] (As shown in Fig. 4) and using the persistence of vision of the human eye, different colors of light can be produced, which will result in the final LED-OFC different in color and color change cycles. To do this, the phase difference between the red, green, blue LEDs is changed, which will make the color features of the final LED-OFC changed. And the specific changes are detailed in the next section.

2.2.2. Generation of LED-OFC with different features

The recognition of CNN is based on the features of LED-OFC, such as the width of stripes, the number of stripes, the color of the stripes and the color change cycle, while the features are selected automatically by CNN. And to produce various LED-OFCs with different features, three kinds of parameters are considered, including frequency, phase difference and distance, as shown in Figs. 5, 6, 7.

By increasing the frequency, the width of the stripe will decrease, and the number of stripes will increase. An example is shown in Fig. 5, the frequency of LED-OFC in Fig. 5(a) is 500 Hz, which is half of Fig. 5(b). And other parameters are same, hence the number of stripes in Fig. 5(a) is half of Fig. 5(b), and the width of the stripe in Fig. 5(a) is twice that in Fig. 5(b), while the duty-ratio of both LED-OFCs in Fig. 5(a), (b) is 50%. As we can see, the higher the frequency, the narrower the width of the stripe, and the more the number of stripes.

By changing the phase difference between different LEDs, the number and the width of the stripe, the color of the stripe and the color change cycle of the final LED-OFC can all be changed. Fig. 6 shows an example, the multi-channel image is the final LED-OFC image, while the single channel image indicates the R, G, B channel separated image of the final LED-OFC image. As identified by the dotted lines in the single channel images, the phase difference between different LEDs in Fig. 6(a) is all 90°, while it is 135° in Fig. 6(b). Other parameters are same, the frequency is 500 Hz, and the distance is 10 cm. In Fig. 6(a), the phase difference between single channel images is smaller than the phase difference in Fig. 6(b). By comparing the final LED-OFC images, we can see that the features such as color of the stripe and color change cycle are very different, which is good for subsequent recognition.

In addition, the distance between LED and CMOS camera is also an influencing factor, but different from the above two parameters, the recognition results should be the same at different distances if the above two parameters are the same. So it is taken out separately to avoid

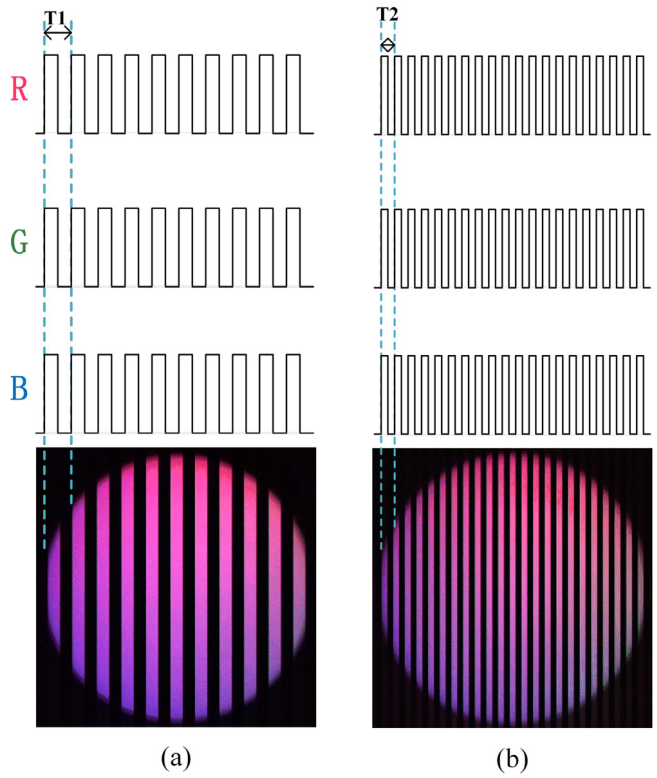

Fig. 5. The effect of different frequencies. (a) LED-OFC of frequency 500 Hz; (b) LED-OFC of frequency 1000 Hz.

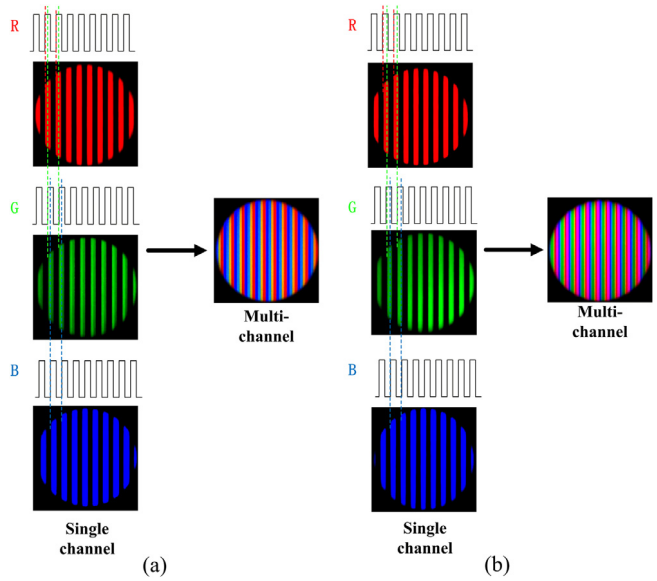

Fig. 6. The effect of different phase difference. (a) LED-OFC with phase difference of 90°; (b) LED-OFC with phase difference of 135°.

misunderstanding. As the distance increases, the number of stripes decreases, but other features do not change. And an example is shown in Fig. 7, In Fig. 7(a), the distance is 20 cm, while in Fig. 7(b), the distance is 40 cm, and other parameters are: the frequency is 500 Hz, and the phase difference is 0°. As we can see, the number of stripes in Fig. 7(a) is twice that in Fig. 7(b), while the width of the stripe and the duty-ratio are the same for both LED-OFCs. For VLC based on CMOS sensor using traditional modulation and demodulation method, the data frame will be partially lost if the distance is large, since when

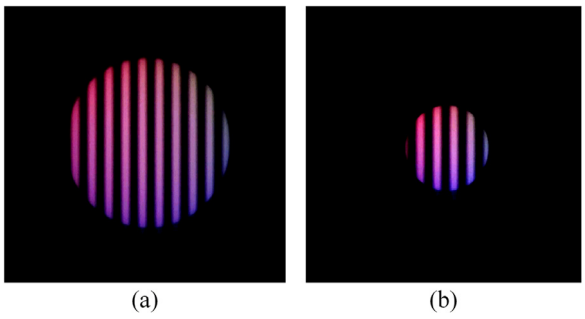

Fig. 7. The effect of the distance between LED and CMOS camera. (a) 20 cm; (b) 40 cm.

the number of stripes decreases to a certain extent, it is not enough to represent a complete data frame by the LED-OFC. This is the reason why traditional methods cannot achieve long-distance recognition.

2.3. LED-OFC recognition

CNN has been very successful in solving computer vision (CV) problems [27–29]. And as mentioned above, by modulating the RGB-LED with different parameters, the LED-OFCs captured by the CMOS sensor can have different features. This converts the problem into a recognition problem, that is, the main work of is to classify and recognize different types of LED-OFC which is what CNN is good at, and then the related information can be obtained by querying the LED-OFC-INFO database. At the same time, this also realizes the combination of VLC and CNN and finally achieves high accuracy and robust recognition of the LED-OFC.

The most common form of a CNN architecture stacks a few convolutional-activation layers, follows them with pooling layers, this pattern is repeated till the image size has been merged spatially small. Then, fully connected (FC) layers are used to map the learned data distribution to the sample space, and the last FC layer holds the output, such as the class scores. The common pattern is shown in Fig. 8 (the input LED-OFC image is cropped manually for demonstration purpose).

However, the situation has changed in the last 2 or 3 years. A new form of CNN architecture called "full convolutional network" (FCN) has been widely used [30], which replaces the FC layers with convolution (CONV) layers. As shown in Fig. 9 (the input LED-OFC image is cropped manually for demonstration purpose), the number of convolutional kernels of last CONV layer is the same as the number of output classes, and the last CONV layer just holds the class scores. This style can not only reduce the number of parameters, but also use the context of the image better [30]. And practice shows that the FCN style usually gets better results. So, a CNN architecture with FCN style is employed in this paper.

In addition, we briefly introduce the mathematical principles of CNN as follows.

2.3.1. Convolutional layer (CONV layer)

The CONV layer is the core of CNN. For a certain layer, the two-dimensional convolutional operation can be described as follows:

$$y(i,j) = \sum_{w}\sum_{h} x(i-w, j-h)\, f(w,h). \tag{3}$$

In Eq. (3), x represents the input with pixel coordinates i and j, f is a convolutional kernel in a CONV layer with pixel coordinates w and h, and y is the output with pixel coordinates i and j. The operation is shown in Fig. 10:

The size of convolutional kernel is 2×2, and the size of the input image is 3×3, and the letters in the small squares in the above figure represent the value of each pixel respectively. The convolution kernel slides from left to right and from top to bottom on the input image, and finally the output is obtained.

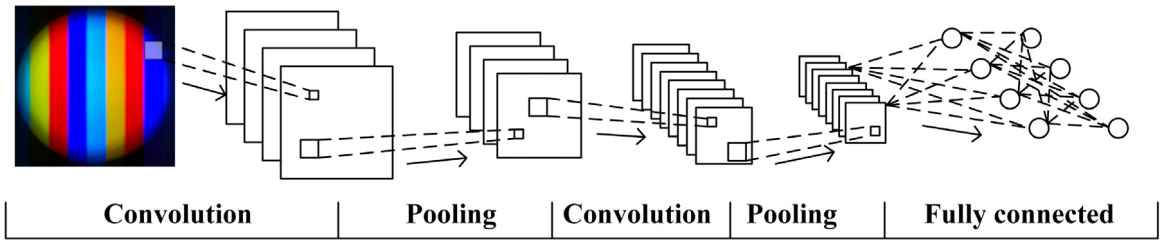

Fig. 8. A typical pattern of convolutional neural network.

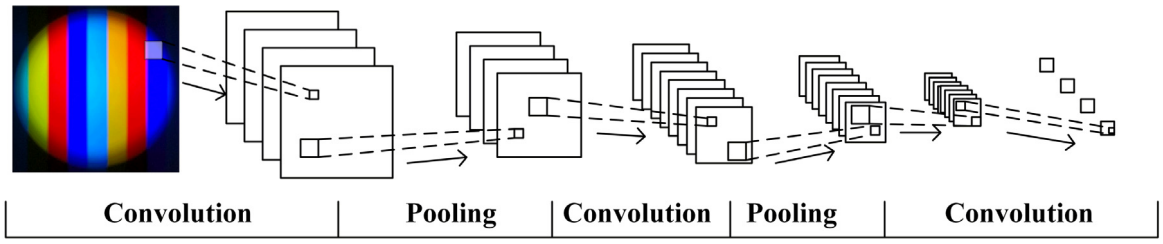

Fig. 9. A typical pattern of the full convolutional network.

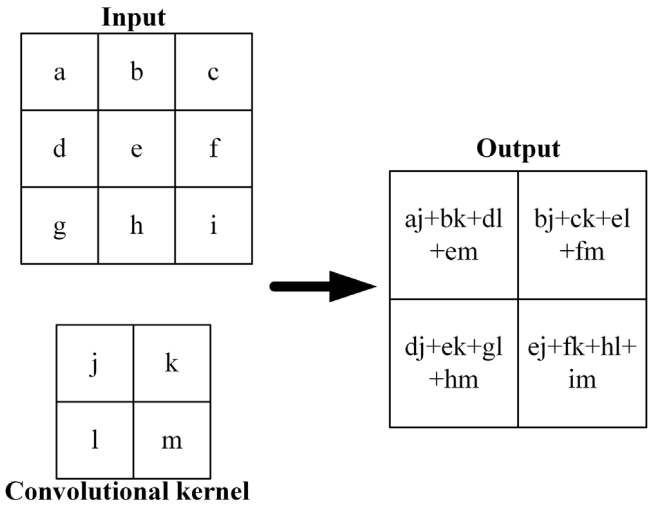

Fig. 10. Convolution operation.

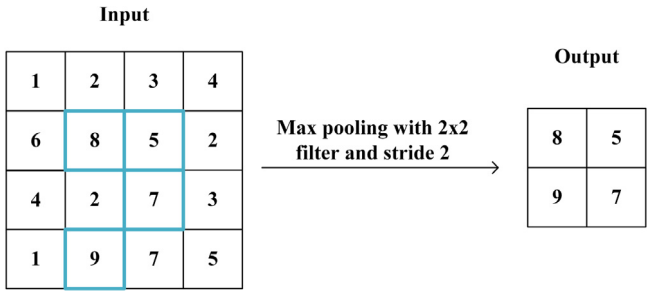

Fig. 11. The illustration of max pooling.

2.3.2. Activation layer

The powerful nonlinear fitting ability of CNN stems from the activation layer, and an activation layer is usually added after the CONV layer. The activation can be described in Eq. (4):

$$a(i,j) = F(y(i,j) + b). \tag{4}$$

In Eq. (4), y represents for output before activation with pixel coordinates i and j, a represents for activated output with pixel coordinates i and j, b represents for bias term, and F is the activation function. There are some often used activation functions, for example, the sigmoid function:

$$F(x) = \frac{1}{1+e^{-x}}. \tag{5}$$

However, the sigmoid function may cause the vanishing gradient problem [31]. Another alternative activation function is the rectified linear unit (RELU), which is widely used in CNN:

$$F(x) = \max(0, x). \tag{6}$$

RELU is the most commonly used activation function, and many experiments have proved its effectiveness. So, in this paper, we use the RELU as the activation function.

2.3.3. Pooling layer

A pooling layer is often used between two CONV layers, which is used to reduce the number of parameters, and hence to control overfitting. And there are two modes: one called max pooling, which takes the maximum value in the pooling region; the other called average pooling, which takes the average value of the pooling region. And practice shows that the max pooling can usually get a better train result, so in this paper, we use max pooling, which is shown in Fig. 11.

2.3.4. Loss function

The loss function is the metric to measure the network's performance, which enables us to optimize the parameters of the network at the same time. There are some often used loss functions, such as square loss (MSE), log loss and hinge loss. Since the LED-OFC recognition belongs to classification field, we use the cross-entropy loss function (belongs to log loss) to measure the similarity between prediction probability distribution and true probability distribution, as shown in Eq. (7):

$$L\left(y_{truth}, y_{prediction}\right) = -\sum y_{truth} log\left(y_{prediction}\right). \tag{7}$$

In Eq. (7), y_{truth} represents the true probability distribution, while $y_{prediction}$ is the predicted probability distribution. As the loss value decreases, the similarity between prediction probability distribution and true probability distribution increases, which means the performance of the network is better.

In addition, to transform the final output into probability distribution, a SoftMax layer is employed in the paper, as Eq. (8) shows:

$$p\left(y_j\right) = \frac{e^{y_j}}{\sum_k e^{y_k}}. \tag{8}$$

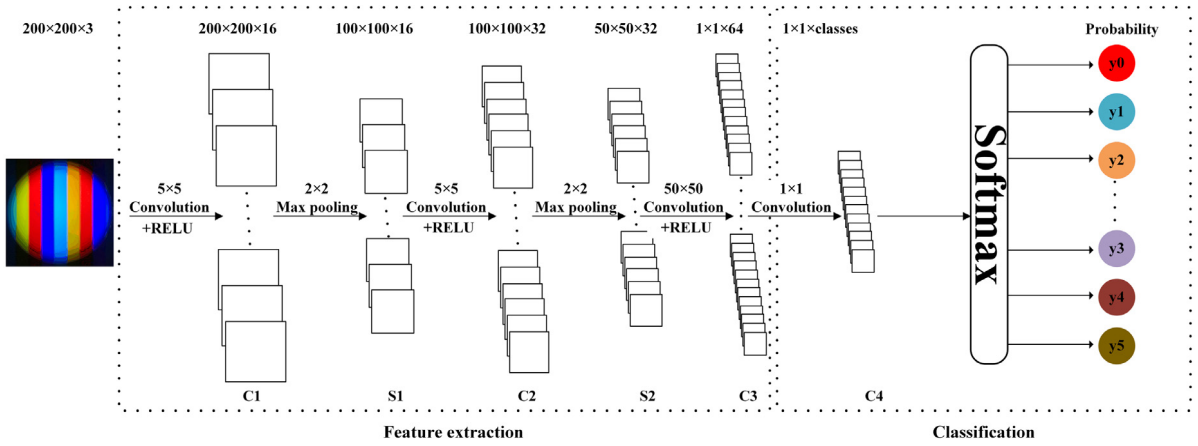

Fig. 12. The architecture of the used CNN.

In Eq. (8), e^{y_j} is the normalized score of the jth class, $\sum_k e^{y_k}$ is the sum of all normalized scores, and the result just represents the probability of each class.

In our system, a CNN architecture which is similar with the "LeNet-5" in [31] is used, while the FC layer is replaced by the CONV layer, as shown in Fig. 12 (the input LED-OFC image is cropped manually for demonstration purpose).

The input of the image is 200×200 in pixels, and the CNN architecture contains seven layers, including three CONV-RELU layers C1, C2 and C3, two max pooling layers S1 and S2, CONV layer C4 and a SoftMax layer. Every layer contains a series of planes, which are referred to as feature maps. Layer C1 contains 16 feature maps of 200 × 200 in pixels, after the first max pooling, the number of feature maps in layer S1 becomes 32 of 100 × 100 in pixels. This pattern is repeated until layer C3, which has 64 feature maps of 1 × 1 in pixels. A CONV layer C4 is added to convert learned features into scores for each class. Finally, a SoftMax layer is used to map the final scores to the (0,1) interval, which indicates the probability that the prediction results belong to each class. The network is trained using TensorFlow [32] and Adam optimizer [33] is used for optimization. And a 12 GB NVIDIA GTX 1080Ti GPU is used for training and test.

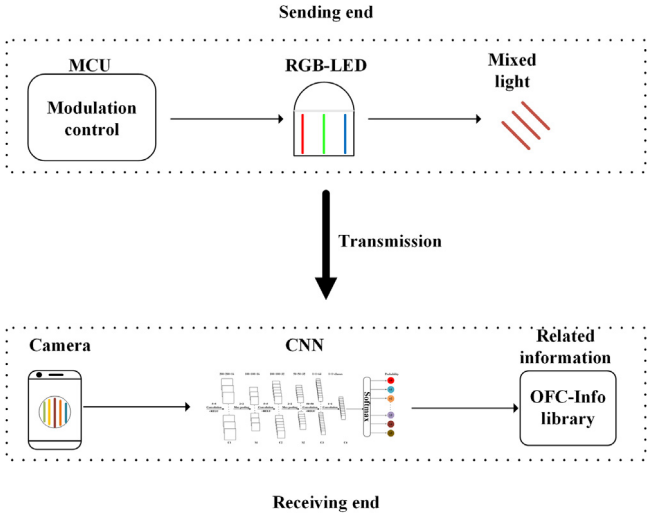

Fig. 13. System overview.

3. Experiments

3.1. Experiments setup

We now provide a brief overview of our system. The RGB-LED is modulated by the micro controller unit (MCU) to produce mixed light. Specially, the RGB-LED is modulated with different distance, frequency and phase difference parameters and that will make the light color mixed. A CMOS camera in the smartphone is used as receiver and rolling shutter mechanism is exploited, which will make the resulting LED-OFC captured by the CMOS camera different in color, stripe width, number of stripes. Then the CNN architecture mentioned before is used to recognize the LED-OFC with different features. Finally, the related information of the recognized LED-OFC can be obtained. The schematic diagram of the system is shown in Fig. 13.

The key hardware parameters are shown in Table 1, and the experiment platform is shown in Fig. 14.

To prove the feasibility of the proposed scheme, a series of experiments are demonstrated. For demonstration purposes, in this paper, different LED-OFCs are given different IDs as the related information, which are used to measure the recognition accuracy.

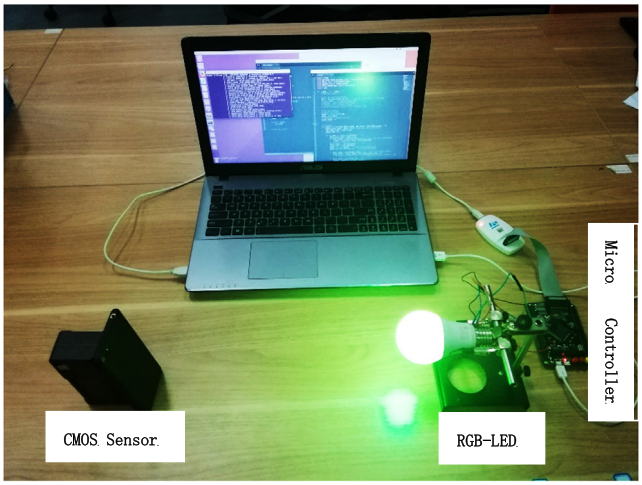

Fig. 14. The overview of experiment platform.

Table 1
Hardware parameters of the experiments.

Parameter	Value
The focal length/mm	4.25
The resolution of the camera	4032 × 2448
The exposure time of the camera/ms	0.01
The ISO of the camera	100
The aperture of the camera	F1.7
The diameter of the LED downlight/cm	6
The power of each LED/W	9
Current of each LED/mA	85
Voltage of each LED/V	5

3.2. Experiments analysis

3.2.1. Frequency resolution and LED-OFC recognition result analysis

As the frequency changes, the number and width of the stripes in the LED-OFC also changes. The comparison of LED-OFC with different frequencies is shown in Fig. 15 (the LED-OFC image is cropped manually for demonstration purpose), and other parameters are: the phase difference is all 0°, and the distance is 10 cm. It is hard to distinguish two different LED-OFCs if the frequency difference of the two LED-OFCs is too small, because the variation in the number and width of the stripe is too small, so there is a threshold. In addition, as mentioned before, the minimum frequency difference is different between the high frequency range and the low frequency range, since in the high frequency range, the stripes are denser, and the stripe width variation due to frequency changes is more difficult to distinguish. In this paper, the low frequency range is defined as 500 Hz to 5000 Hz, while the high frequency range is from 5000 Hz to 10000 Hz. And two experiments in the low frequency range and the high frequency range are conducted separately. Specifically, the impact of different frequency resolutions on LED-OFC recognition (Frequency resolution indicates the frequency difference of LED-OFCs with different frequencies) is explored. Other parameters are: the phase difference is 0°, the distance is 20 cm. And for each frequency resolution, about 2000 images are acquired for training, and 200 images for test. The results are shown as follows.

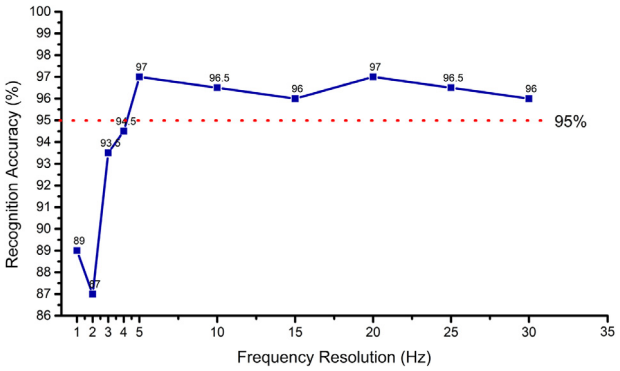

Fig. 16. The recognition accuracy rate at different frequency resolution in low resolution range.

3.2.1.1. Low frequency range. As Fig. 16 shows, in the low frequency range, recognition accuracy increases rapidly with increasing frequency resolution and tends to stabilize quickly. When the frequency resolution is lower than 3 Hz, the LED-OFC recognition accuracy rate drops a lot. And the LED-OFC recognition accuracy rate is greater than 95% when the frequency resolution is higher than 5 Hz. Hence, to ensure high recognition accuracy rate and robustness, the frequency resolution in the low frequency range should higher than 5 Hz.

3.2.1.2. High frequency range. As Fig. 17 shows, in high frequency range, the overall accuracy rate increases more slowly with increasing frequency resolution compared with that in the low frequency range. The minimum frequency resolution is 40 Hz, which is larger compared with the minimum frequency resolution in the low frequency range. This is also in line with our expectation that the minimum frequency resolution in the high frequency range will be higher due to the excessively dense stripes of the LED-OFC.

Combine the above two experiments, we can draw conclusions that, changes in the number and width of stripes caused by frequency changes can be used as a basis for CNN recognition. And to make the

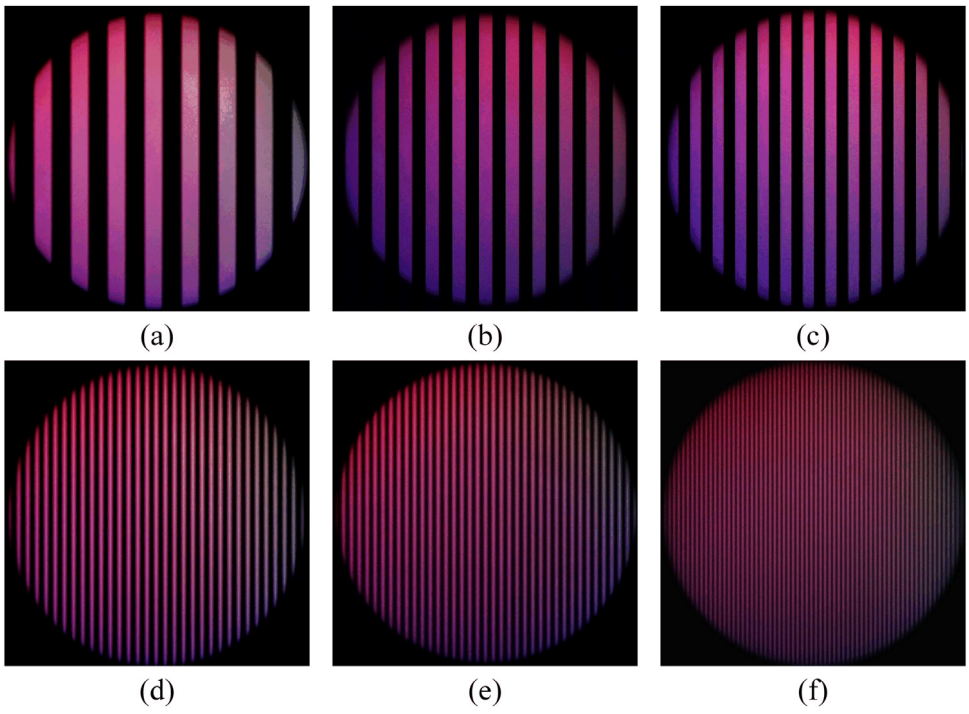

Fig. 15. LED-OFC of different frequencies. (a) 200 Hz; (b) 500 Hz; (c) 600 Hz; (d) 800 Hz; (e) 1000 Hz; (f) 1500 Hz.

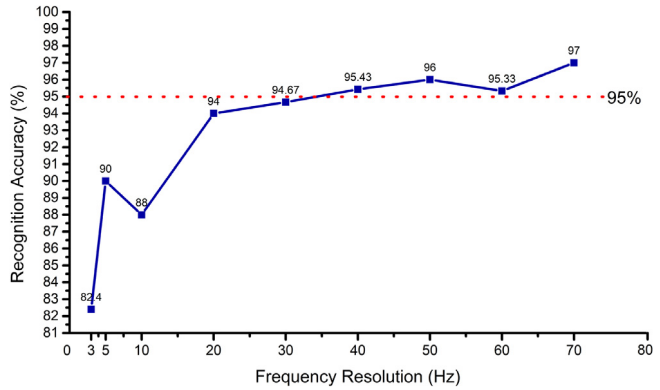

Fig. 17. The recognition accuracy rate at different frequency resolution in high frequency range.

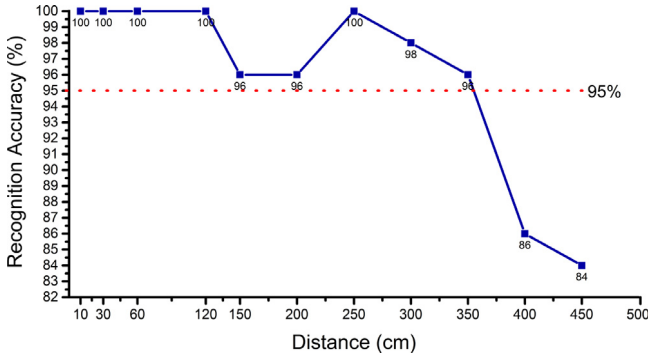

Fig. 19. Recognition accuracy rate at different distance.

LED-OFC recognition with high accuracy rate, the minimum frequency resolution is 40 Hz. Furthermore, the minimum frequency resolution is still lower than that in [14] even in high resolution range, which only achieves a minimum frequency resolution of 60 Hz. And considering that the author does not experiment separately in the high and low frequency range, which would cause a large error, the result is good enough. The frequency range used in this experiment is from 500 Hz to 10000 Hz, which could provide 1025 ($\frac{(10000-5000)}{40}+\frac{(5000-500)}{5}=1025$) different LED-OFCs. Moreover, here we only consider the case that the frequency of the three LEDs is the same, if not, the types of LED-OFC will be much more according to the arrangement and combination, since the color of the stripes will change, and the color feature will be a basis for CNN to recognize, including the color and the color change cycle. And the rough number can be calculated as follows: There are 1025 types of frequency, as the result shows. And for the RGB-LED, the frequency of each LED is independent, hence there are 3 different frequencies. According to the principle of arrangement and combination, the final amount should be: 10^9 (A^3_{1025}).

3.2.2. *Distance and LED-OFC recognition result analysis*

As distance increases, the light intensity received by CMOS camera drops, the number of stripes in the LED-OFC image also decreases, then the recognition difficulty increases. The comparison of LED-OFCs with different distances is shown in Fig. 18 (the LED-OFC image is cropped manually for demonstration purpose), and other parameters are: the phase difference is all 0°, and the frequency is 1000 Hz. To quantify the effect of distance, an experiment is designed, which explores the effect of different distances on recognition accuracy rate. And it is important to note that the focal length is not adjusted when taking pictures of different distances, so for a specific type of LED-OFC, the absolute width of the stripes at different distances does not change, which is different from traditional modulation and demodulation method. Other parameters are as follows: The frequency is from 500 Hz to 9000 Hz with interval 500 Hz (which is enough according to previous experiments) and the phase difference is 0°. And for each distance, about 2000 images are used for training and 200 images are used for test. The result is shown in Fig. 19.

As we can see, the recognition accuracy rate remains high (above 95%) until the distance is more than 3.5 m. And that is shorter than the maximum recognition distance in [14], which achieves maximum

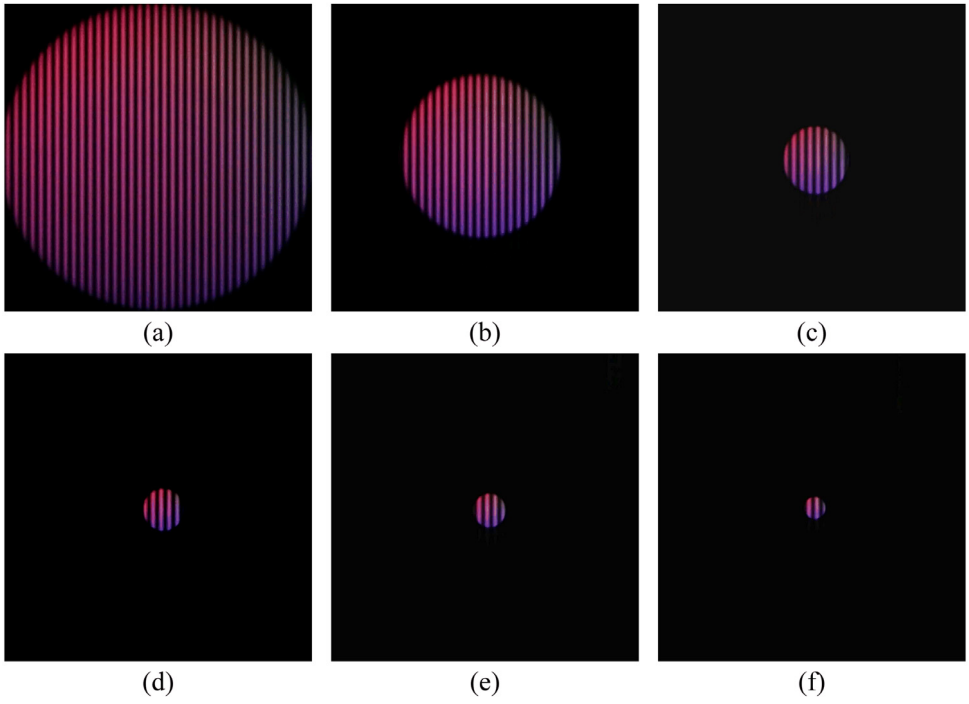

Fig. 18. LED-OFC of different distances. (a) 10 cm; (b) 20 cm; (c) 50 cm; (d) 80 cm; (e) 100 cm; (f) 150 cm.

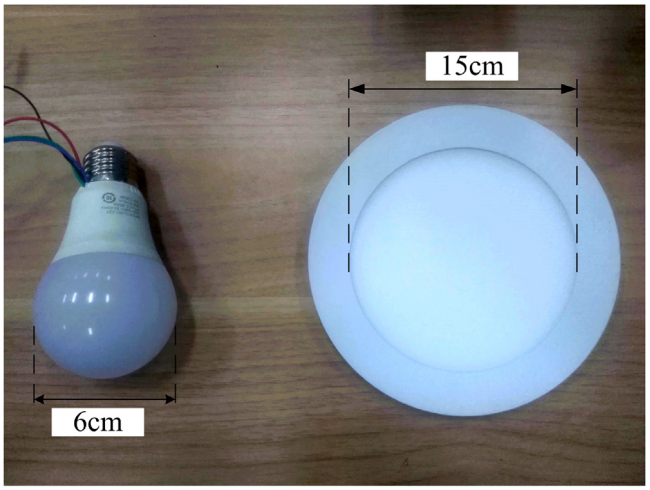

Fig. 20. Comparison of two types of LED.

recognition distance of 6 m. However, considering the transmitter of this experiment only uses LED with a diameter of 6 cm, while in [14] the diameter of the LED is 15 cm, as shown in Fig. 20, the result is acceptable. As the diameter of the LED increases, the diameter of the LED-OFC increases, and the number of stripes also increases, which allows the recognition distance to be further increased. If the diameter of the LED reaches 15 cm, the maximum recognition distance could reach 10 m to 20 m at least. And the maximum recognition distance is a lot longer than recent related works using traditional modulation and demodulation method [16,18–21], which only achieves effective transmission distance of 10 cm to 50 cm, and more importantly, the accuracy rate does not decrease. For the related studies using traditional modulation and demodulation method, the reduction in the number of stripes is unacceptable. For example, most of the recent related research uses the OOK scheme to modulate LED and here we assume the encoding sequence is 16 bits. To transmit the 16 bit data, the number of stripes in the image captured by the CMOS sensor cannot less than 16 (Including bright and dark stripes), otherwise it will cause the loss of data frames. So, the distance is severely restricted in order to ensure the number of stripes. Though the number of stripes can be increased by increasing the modulation frequency, but this will further increase the difficulty of demodulation, which is proved in the frequency resolution experiments before. But for the LED-OFC, there is no such restriction, and the above distance experiment result proves this. Moreover, the above experiment result also proves that the proposed method does not require additional synchronization, since the test images are captured at any time and no additional synchronization with the transmitter beforehand is required. Hence the proposed scheme is more applicability compared with all recent related works [14,16,18–21].

3.2.3. Phase difference resolution and LED-OFC recognition result analysis

The number, the color and the width of stripes can all be changed by changing the phase difference between different LEDs, and some examples are shown in Fig. 21 (the LED-OFC image is cropped manually for demonstration purpose), other parameters are: the distance is 10 cm, and the frequency is 1000 Hz. However, it is difficult to recognize if the difference in phase difference of different LED-OFCs is small. So similar with the experiment of exploring frequency resolution, the impact of different phase difference resolutions on the LED-OFC recognition is explored. Other parameters are shown as follows: the frequency is 1000 Hz, the distance is 10 cm. Same as above, for each phase difference resolution, 2000 images are acquired for training, while 200 images for test. The result is shown in Fig. 22.

The recognition accuracy increases quickly with increasing phase difference resolution when the difference in phase difference is lower than 7.2° and becomes steady afterwards, and when the phase difference between LEDs is greater than 7.2°, the recognition accuracy is greater than 95%, as seen in Fig. 22. This indicates that the color feature of the LED-OFC can be used as a basis for recognizing different kinds of LED-OFCs. And to ensure high recognition accuracy, the phase difference resolution should higher than 7.2°, the number of distinguished LED-OFC type is 50 ($\frac{360}{7.2} = 50$). The types of LED-OFC will be much more if the phase difference between the three LEDs is different, since the color and the color change cycle will be more different. Here

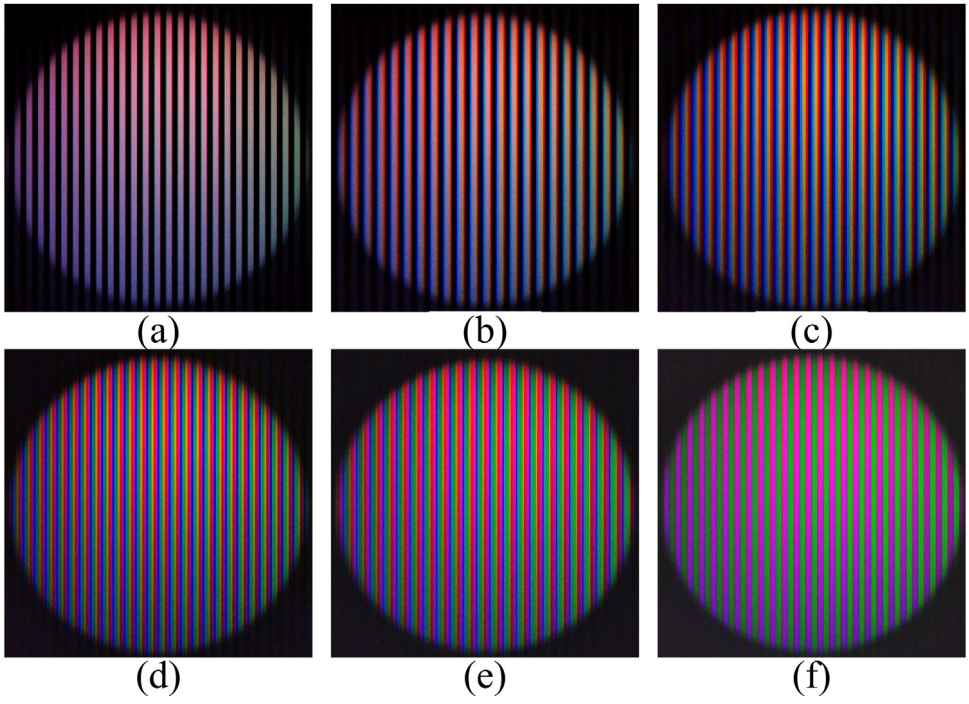

Fig. 21. LED-OFC of different phase difference. (a) 3.6°; (b) 36°; (c) 72°; (d) 108°; (e) 144°; (f) 180°.

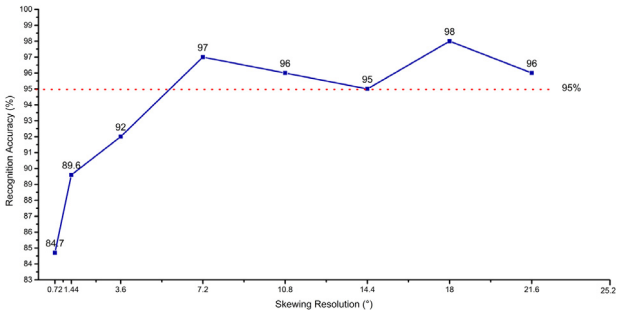

Fig. 22. The recognition accuracy rate at different phase difference resolution.

Table 2
The parameters of different LED-OFCs.

ID	Frequency (Hz)	Phase difference (°)	Distance (cm)
0	500	90	20
1	1000	90	40
2	2000	90	60
3	3000	90	80
4	4000	90	100
5	5000	90	120
6	6000	90	140
7	7000	90	160
8	8000	90	180
9	9000	90	200
10	1000	10.8	20
11	1000	21.6	40
12	1000	43.2	60
13	1000	57.6	80
14	1000	72	100
15	1000	108	120
16	1000	129.6	140
17	1000	144	160
18	1000	162	180
19	1000	180	200

we just consider the case that the phase difference between different LEDs is the same. And the rough number can be calculated as follows: There are 50 types of phase difference, as the result shows. And for the RGB-LED, there is 2 independent phase difference. According to the principle of arrangement and combination, the final amount should be: 10^3 (A_{50}^2).

3.3. Extended experiment and analysis

To simulate the real environment, an extended experiment is designed to explore the combined effects of above parameters. 20 different LED-OFCs are used, and the specific parameters for each LED-OFC are shown in Table 2. And for each LED-OFC ID about 200 images is acquired for training, and 20 images for test. The result is shown in Fig. 23.

For the extended experiment, all the parameters are considered, including distance, phase difference and frequency. For the top 10 LED-OFC IDs, the comprehensive impact of distance and frequency is explored, while for the last 10 LED-OFC IDs, the comprehensive impact of distance and phase difference is explored. And finally, combining all the LED-OFC IDs, the comprehensive impact of the three parameters can be explored. And the result shows that most of the LED-OFCs achieve 100% recognition accuracy, as seen from above. And this also shows from an experimental point that the proposed scheme is highly practical and applicable. And as mentioned before, we only consider the case that the frequency and the phase difference of the three LEDs are the same, the types of LED-OFC will be much more if the frequency and the phase difference of the three LEDs are different, which will further improve practicality.

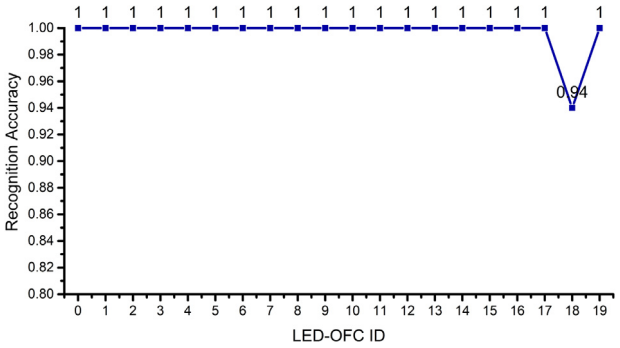

Fig. 23. The result of extended experiment.

4. Conclusion

In this paper, an RGB-LED is introduced and modulated to assign different features to the LED-OFC captured by the CMOS sensor, and an LED-OFC modulation and recognition algorithm based on visible light communication using CNN is proposed to make up the shortcomings of recent VLC based on CMOS sensor which uses traditional modulation and demodulation method, or modulation and recognition method. Specifically, for the modulation and demodulation method, the maximum transmission distance is short, and it is difficult to ensure the synchronization of captured images by the CMOS image sensor. Furthermore, many studies are too focused on real-time performance, which is difficult to achieve at this stage, and for many application scenarios high real-time performance is not required but recognition accuracy is more important. For the modulation and recognition method, demodulation is not required, recognition is enough. However, for current related work, though the maximum recognition distance is larger, the recognition number is far from enough, and the robustness and accuracy are not high under some conditions. This is the first time CNN is applied to the field of VLC based on CMOS sensor and achieves state-of-the-art results. The proposed scheme has no synchronization requirements, and can extend the maximum recognition distance, increase the recognition amount, recognition accuracy and the robustness of the system. The RGB-LED is modulated by different phases which will result in phase difference between different LEDs. And benefit from the rolling shutter mechanism of the CMOS sensor, the LED-OFC captured by CMOS sensor will have stripes with different colors, which greatly enriched the features of the LED-OFC. In addition to phase difference, frequency and distance are also considered, which makes the resulting LED-OFC captured by the CMOS sensor different in color, stripe width, number of stripes. Furthermore, a CNN architecture with FCN style is employed after the LED-OFC is captured, and the LED-OFC of different features can be recognized well by the CNN.

The experiment results verify the superiority of the proposed scheme. As the results show, the maximum recognition distance and recognition accuracy are greatly improved compared with related works [14,16,18–21]. Specially, by the proposed scheme, about 50 thousand ($50 \times 1025 = 51250$) of different LED-OFCs can be recognized with high accuracy (above 95%), which is about 50 times of that in [14]. And this is far lower than the actual quantity, since each LED can be controlled by different parameters and the rough number can reach 10^{12} ($A_{1035}^3 \times A_{50}^2$) as deduced and calculated above, which is enough for most application scenarios. In addition, the maximum recognition distance reaches 3.5 m, which sis far more than all the related research. Therefore, the proposed scheme can be applied to a variety of scenarios.

Acknowledgments

This work supported by National Undergraduate Innovative and Entrepreneurial Training Program, China (No. 201510561003,

Table A.1
The accuracy and loss in low frequency range.

Frequency resolution (Hz)	Accuracy (%)	Loss
1	98.71	0.0779
2	100	0.0262
3	99.13	0.0340
4	99.73	0.0322
5	98.62	0.0263
10	99.17	0.0268
15	100	0.0253
20	100	0.0106
25	100	0.0158
30	100	0.0033

Table A.2
The accuracy and loss in high frequency range.

Frequency resolution (Hz)	Accuracy (%)	Loss
3	97.76	0.1722
5	91.67	0.2567
10	98.00	0.1087
20	100	0.6908
30	97.57	0.0834
40	98.00	0.0742
50	100	0.0734
60	98.36	0.0677
70	98.65	0.0911

Table A.3
The accuracy and loss in different distance.

Distance (cm)	Accuracy (%)	Loss
10	100	6.756e−4
30	100	1.339e−3
60	100	1.005e−3
90	100	0.0144
120	100	0.0150
150	100	7.862e−3
200	100	0.0437
250	100	5.422e−3
300	100	2.302e−3
350	100	0.0263
400	98.13	0.1276
450	93.47	0.5766

201610561065, 201610561068, 201710561006, 201710561054, 201710561057, 201710561058, 201710561199, 201710561202, 201810561217, 201810561195, 201810561218, 201810561219), Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation, China ("Climbing Program" Special Funds) (pdjh2017b0040, pdjha0028), Guangdong science and technology project, China (2017B010114001), National Natural Science Foundation of China, China (No. 11771152), Major Science and Technology Projects of Guangdong Province, China (No. 2015B010128008, 2015B010109006, 2016B010127003).

Appendix

A.1. The training accuracy and loss

See Tables A.1–A.4.

A.2. The validation accuracy and loss

See Tables A.5–A.8.

Table A.4
The accuracy and loss in different phase difference.

Phase difference (°)	Accuracy (%)	Loss
1	94.30	0.1461
3	99.10	0.0412
5	100	0.0110
10	100	0.0199
15	100	0.0163
20	100	0.0018
25	100	0.0036
30	100	0.0025

Table A.5
The accuracy and loss in low frequency range.

Frequency resolution (Hz)	Accuracy (%)	Loss
1	83.33	0.2200
2	91.67	0.0922
3	94.96	0.0961
4	93.67	0.1330
5	94.93	0.2130
10	95.35	0.0421
15	99.56	0.0128
20	99.33	0.0491
25	97.28	0.0208
30	98.72	0.0417

Table A.6
The accuracy and loss in high frequency range.

Frequency resolution (Hz)	Accuracy (%)	Loss
3	91.01	0.2756
5	87.52	0.3812
10	89.00	0.3690
20	96.65	0.1573
30	98.13	0.1303
40	98.33	0.1177
50	99.10	0.0700
60	96.52	0.1144
70	97.29	0.0506

Table A.7
The accuracy and loss in different distance.

Distance (cm)	Accuracy (%)	Loss
10	100	1.162e−3
30	100	1.069e−3
60	100	7.253e−3
90	100	0.0549
120	100	0.0375
150	94.87	0.0759
200	97.90	0.0567
250	100	9.623e−3
300	100	4.307e−3
350	95.80	0.1609
400	88.05	0.3432
450	86.32	0.5877

Table A.8
The accuracy and loss in different phase difference.

Phase difference (°)	Accuracy (%)	Loss
1	83.40	0.6440
3	88.40	0.3260
5	93.50	0.1361
10	95.33	0.1527
15	97.10	0.0513
20	100	0.0014
25	100	0.0020
30	100	0.0011

References

[1] A. Jovicic, J. Li, T. Richardson, Visible light communication: opportunities, challenges and the path to market, IEEE Commun. Mag. 51 (12) (2013) 26–32.

[2] D. Karunatilaka, F. Zafar, V. Kalavally, R. Parthiban, LED based indoor visible light communications: State of the art, IEEE Commun. Surv. Tutor. 17 (3) (2015) 1649–1678.
[3] L. Feng, R.Q. Hu, J. Wang, P. Xu, Y. Qian, Applying VLC in 5G networks: Architectures and key technologies, IEEE Network 30 (6) (2016) 77–83.
[4] P.F. Mmbaga, J. Thompson, H. Haas, Performance analysis of indoor diffuse VLC MIMO channels using angular diversity detectors, J. Lightwave Technol. 34 (4) (2016) 1254–1266.
[5] R. Mitra, V. Bhatia, Chebyshev polynomial-based adaptive predistorter for nonlinear LED compensation in VLC, IEEE Photonics Technol. Lett. 28 (10) (2016) 1053–1056.
[6] H.H. Lu, C.Y. Li, C.A. Chu, T.C. Lu, B.R. Chen, C.J. Wu, D.H. Lin, 10 m/25 Gbps LiFi transmission system based on a two-stage injection-locked 680 nm VCSEL transmitter, Opt. Lett. 40 (19) (2015) 4563–4566.
[7] Ifthekhar, M.d. Shareef, et al., Neural network-based indoor positioning using virtual projective invariants, Wirel. Pers. Commun. 86 (4) (2016) 1813–1828.
[8] M.S. Ifthekhar, N.T. Le, M.A. Hossain, T. Nguyen, Y.M. Jang, Neural network-based indoor positioning using virtual projective invariants, Wirel. Pers. Commun. 86 (4) (2016) 1813–1828.
[9] D. Zheng, G. Chen, J.A. Farrell, Joint measurement and trajectory recovery in visible light communication, IEEE Trans. Control Syst. Technol. 25 (1) (2017) 247–261.
[10] Y.S. Eroğlu, Güvenç. I, A. Şahin, .Y. Yapıcı, N. Pala, M. Yüksel, Multi-element VLC networks: LED assignment, power control, and optimum combining, IEEE J. Sel. Areas Commun. 36 (1) (2018) 121–135.
[11] H. Steendam, A 3-D positioning algorithm for AOA-based VLP with an aperture-based receiver, IEEE J. Sel. Areas Commun. 36 (1) (2018) 23–33.
[12] C.W. Chow, C.Y. Chen, S.H. Chen, Visible light communication using mobile-phone camera with data rate higher than frame rate, Opt. Express 23 (20) (2015) 26080–26085.
[13] Y. Liu, Decoding mobile-phone image sensor rolling shutter effect for visible light communications, Opt. Eng. 55 (1) (2016) 016103.
[14] C. Xie, W. Guan, Y. Wu, L. Fang, Y. Cai, The LED-ID detection and recognition method based on visible light positioning using proximity method, IEEE Photonics J. 10 (2) (2018) 1–16.
[15] W. Guan, Y. Wu, C. Xie, L. Fang, X. Liu, Y. Chen, Performance analysis and enhancement for visible light communication using CMOS sensors, Opt. Commun. 410 (2018) 531–551.
[16] B. Zhang, K. Ren, G. Xing, X. Fu, C. Wang, SBVLC: Secure barcode-based visible light communication for smartphones, IEEE Trans. Mob. Comput. 15 (2) (2016) 432–446.
[17] ISO/IEC 18000-3:2010 Information technology – Radio frequency identification for item management – Part 3: Parameters for air interface communications at 13.56 MHz.
[18] H.W. Chen, S.S. Wen, Y. Liu, M. Fu, Z.C. Weng, M. Zhang, Optical camera communication for mobile payments using an LED panel light, Appl. Opt. 57 (19) (2018) 5288–5294.
[19] C.W. Chen, C.W. Chow, Y. Liu, C.H. Yeh, Efficient demodulation scheme for rolling-shutter-patterning of CMOS image sensor based visible light communications, Opt. Express 25 (20) (2017) 24362–24367.
[20] C.W. Chow, R.J. Shiu, Y.C. Liu, X.L. Liao, K.H. Lin, Y.C. Wang, Y.Y. Chen, Using advertisement light-panel and CMOS image sensor with frequency-shift-keying for visible light communication, Opt. Express 26 (10) (2018) 12530–12535.
[21] C.W. Chow, R.J. Shiu, Y.C. Liu, C.H. Yeh, X.L. Liao, K.H. . . . Lin, Y.Y. Chen, Secure mobile-phone based visible light communications with different noise-ratio light-panel, IEEE Photonics J. 10 (2) (2018) 1–6.
[22] C. Danakis, M. Afgani, G. Povey, I. Underwood, H. Haas, Using a CMOS camera sensor for visible light communication, in: Globecom Workshops (GC Wkshps), 2012 IEEE, IEEE, 2012, pp. 1244–1248.
[23] C.W. Chow, C.Y. Chen, S.H. Chen, Enhancement of signal performance in LED visible light communications using mobile phone camera, IEEE Photonics J. 7 (5) (2015) 1–7.
[24] D.T. Nguyen, Y. Park, Data rate enhancement of optical camera communications by compensating inter-frame gaps, Opt. Commun. 394 (2017) 56–61.
[25] B. Fahs, J. Chellis, M.J. Senneca, A. Chowdhury, S. Ray, A. Mirvakili, P... Zarkesh-Ha, A 6-m OOK VLC link using CMOS-compatible pn photodiode and red LED, IEEE Photon. Technol. Lett. 28 (24) (2016) 2846–2849.
[26] C. CIE, Commission Internationale De L'eclairage Proceedings, 1931, Cambridge University Press Cambridge, 1932.
[27] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[28] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, in: Advances in Neural Information Processing Systems, 2015, pp. 91–99.
[29] J. Redmon, A. Farhadi, YOLO9000: better, faster, stronger. arXiv preprint, 2017.
[30] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
[31] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[32] S.S. Girija, Tensorflow: Large-scale machine learning on heterogeneous distributed systems, 2016.
[33] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.