A Deep Learning-Assisted Visible Light Positioning Scheme for Vehicles With Image Sensor

Jing He and Biao Zhou

***Abstract*—Visible light communication (VLC) based on image sensor (IS) as the receiver is one of the supplementary technologies of radio frequency communication. By combining with image processing technology, VLC has various applications in smart home, underwater communication and intelligent transportation system (ITS). In the paper, it focuses on the vehicle positioning service provided by VLC in ITS. For the weak image spatial separability caused by long-distance VLC transmission between the vehicle and the infrastructure, it will lead to serious deterioration in communication and positioning performance. Thus, a deep learning-assisted IS-based visible light positioning (VLP) scheme is proposed and experimentally demonstrated. At the transmitter, a novel coding method is proposed to support both short-distance and long-distance communication based on VLC in ITS. Meanwhile, it can meet the dimming requirements of LED road infrastructure. In addition, in the proposed VLP scheme, it uses an artificial neural network (ANN) to predict the location of the vehicle. It is demonstrated that long-distance communication, high-accuracy positioning, and LED dimming can be realized simultaneously. The results show that, by using the proposed scheme, as the transmission distance is 2 m, the bit error rate (BER) of system is 1.25×10^{-4}. And the average positioning error is 19.8 mm at the maximum simulated distance of 30 m.**

***Index Terms*—Visible light communication, vehicle positioning, image sensor, artificial neural network.**

I. Introduction

ACCURATE vehicle positioning is essential to provide services of automatic driving, assisted driving, automatic parking, and trajectory tracking for intelligent transportation systems (ITS). With the development of Internet of Vehicles (IoV), light emitting diodes (LEDs) have become more and more widely used in various road infrastructure and vehicle lighting systems, and the application of image sensor (IS) in vehicles has also developed from driving record to pedestrian detection, lane detection, and vehicle detection to provide safe driving [1]. Thus, the vehicle applications based on visible light communication (VLC) with LEDs and IS are paid more attention [2]. The visual geometry of IS can be used to achieve high-precision positioning [3] and it is considered as one of the effective solutions to provide vehicle positioning services [4].

Manuscript received 25 April 2022; revised 17 June 2022; accepted 28 June 2022. Date of publication 7 July 2022; date of current version 20 July 2022. This work was supported in part by the National Science Foundation of Hunan Province under Grant 2020JJ4210 and in part by the National Natural Science Foundation of China under Grant 61775054. (*Corresponding author: Jing He.*)

The authors are with the College of Computer Science and Electronic Engineering, Hunan University, Changsha 410012, China (e-mail: jhe@hnu.edu.cn).

Digital Object Identifier 10.1109/JPHOT.2022.3188628

Due to the large Field of view (FOV), strong spatial separability and flexible mobility, IS can be used as a detector in VLC. In general, IS has two shutter modes such as rolling shutter and global shutter mode. Rolling shutter can usually achieve higher data rate because its sampling rate is much higher than the frame rate. To improve the BER performance, sampling scheme based on length estimation [5], column matrix selection scheme based on energy diffusion [6], ANN-based decoding scheme [7], and mapping-based demodulation scheme [8] are proposed. For supporting mobility, column selection scheme based on gray values [9] and decoding scheme based on K-means clustering [10] are reported. In addition, CNN-based decoding scheme [11], color and spatial multiplexing [12] and multiple input multiple output (MIMO) technology are proposed to improve the data rate. However, based on the rolling shutter, the communication distance and data rate are determined by the size of the LED light source in the image, the communication distance is limited, and it is usually tens of centimeters. For global shutter, under-sampling phase shift OOK modulation [13], S2-PSK modulation [14], under-sampling frequency shift OOK modulation [15], pulse width/position modulation [16], color shift keying modulation [17] are investigated. It can support long-distance transmission between the vehicle and the infrastructure, but the data rate is low, and it is less than the IS frame rate.

Recently, Global Positioning System (GPS) is the most popular vehicle positioning technology due to its wide coverage and low cost. However, GPS signals are susceptible to multipath effects. It usually gets the lack of stability and reliability in urban, underground, and tunnel environments, and its positioning accuracy is about 10 m [18]. Although LIDAR and RADAR positioning technologies can achieve higher positioning accuracy, high equipment cost is required [19]. Positioning technologies based on radio frequency as WIFI and Bluetooth, are susceptible to electromagnetic interference and co-frequency interference, and the positioning performance in downtown areas will drop sharply due to the increased access from multiple vehicles [20]. The visible light positioning (VLP) based on IS can achieve high positioning accuracy with the centimeter level due to its inherent geometric imaging. Meanwhile, the spatial separation characteristics of IS can effectively separate signal light sources and noise light sources, and it have good anti-noise ability [21].

VLC-based positioning using IS can be applied in both indoor and outdoor environment. For indoor positioning, it is focused on the improvement of positioning accuracy by using the positioning algorithms [22]–[24]. The positioning scheme based on

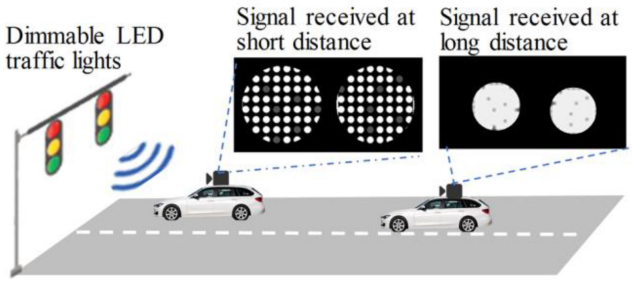

Fig. 1. IS-based VLP for vehicles.

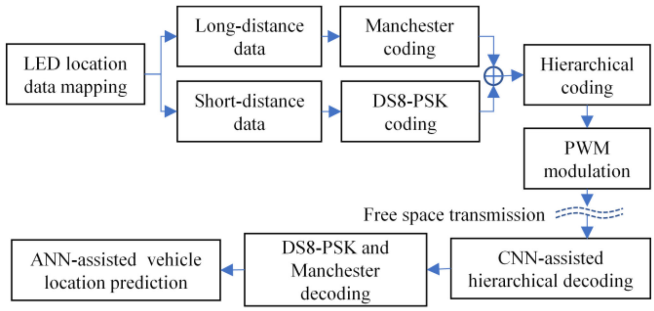

Fig. 2. The proposed deep learning-assisted VLP scheme.

visual scaling factor [22], or based on n-point 3D to 2D mapping [23], and the non-line-of-sight positioning scheme based on K pairwise [24] are proposed. In general, the achieved positioning accuracy is centimeter. However, the data communication range between the transmitter and the receiver is not often considered and the positioning range is limited. In addition, for outdoor environment, IS-based VLP solutions are attracted attention in IoV and it is achieved by using mapping from 3D coordinates to 2D coordinates based on visual geometry [25]–[33]. The positioning scheme based on perspective transformation [25], neural network [26], geometric principles [27], photogrammetry and particle filter [28], collinearity and image artifact elimination positioning scheme [29], vehicle tracking scheme based on perspective transformation [30], positioning performance analysis under the influence of Gaussian white noise [31], distance estimation based on Kalman filter aided Dual geometric vision [32], and Kalman filter aided geometric positioning [33] are proposed. However, the communication performance between the vehicle and the LED source based on VLC is not investigated for most of them. Only spatial 2-phase shift keying (S2-PSK) modulation and on-off keying (OOK) modulation are used to achieve vehicle communication and positioning simultaneously, and the transmission distance is limited [27], [28]. In addition, the performance of IS-based VLP technology is limited by the transmission distance. As shown in Fig. 1, as the transmission distance between the vehicle and the dimmable LED traffic lights increases, the area of the light source captured by the IS becomes smaller, high-frequency components are lost, and the spatial separation of the LEDs decreases. Thus, the communication and positioning performance is degraded.

In the paper, a neural network-assisted IS-based vehicle positioning scheme is proposed and demonstrated. For the proposed scheme, at the transmitter, a novel coding method is proposed to support both short-distance and long-distance transmission between the vehicle and the infrastructure. Meanwhile, it can increase the data rate and meet the dimming requirements of LED road infrastructure by adjusting the LED lighting power based on the ambient light intensity. At the receiver, a convolutional neural network (CNN) based decoder is used and an artificial neural network (ANN) is adopted to realize vehicle positioning. In addition, by using the proposed scheme, the performance of bit error rate (BER) and the average positioning error are investigated.

II. Principle

The architecture of the proposed neural network-assisted IS-based VLP scheme is shown in Fig. 2. Firstly, for the vehicle, as it is far from and near the LED traffic light, the received location data of the LED is different. In this way, at the transmitter, the location data of the LED traffic light is mapped into two different data. One is long-distance data that each bit of it is transmitted by multiple LEDs on the LED matrix in a time slot. And it has strong spatial separability to support long-distance transmission. The other is short-distance data that each bit of it is sent by a single LED on the LED matrix in a time slot. And it has weak spatial separability to support only short-distance transmission. Meanwhile, long-distance data and short-distance data are respectively encoded by Manchester and dimmable spatial eight-phase shift keying (DS8-PSK) to realize the dimming control of the LED traffic light. Secondly, the two data are hierarchical coded for long- and short-distance transmission simultaneously. Thirdly, the hierarchical coded data is modulated with pulse width modulation (PWM) and transmitted by the LED light. After transmission over free space, it is received by the IS installed on the vehicle. The received signal is decoded by CNN to get the layered data, and then the position information of the LED street light is obtained by the decoding of DS8-PSK and Manchester. Finally, the position of the vehicle is output through an ANN.

A. Transmitter

At the transmitter, an 8×8 LED matrix is used. To realize the dimming of the LEDs for short-distance transmission, the short-distance data is encoded by DS8-PSK [34]. The DS8-PSK signal consists of a LED reference group and LED data groups. For the 8×8 LED matrix, one 2×4 LED matrix is the reference group, and seven 2×4 LED matrix are the data groups. In the paper, the phase value of the LED reference group is set as 0.

Fig. 3 shows the DS8-PSK signal of an LED data group with 8 LEDs such as LED #1 …LED #8 under the dimming level of 7/8. Every three bits in the short-distance data is regarded as a cluster, and then mapped to a phase shift value according to Table I. Subsequently, the three bits of the cluster are modulated to an LED data group, and it is based on the difference between the phase shift value and the phase value of the LED reference group, as shown in Table II. In Table II, it shows the mapping

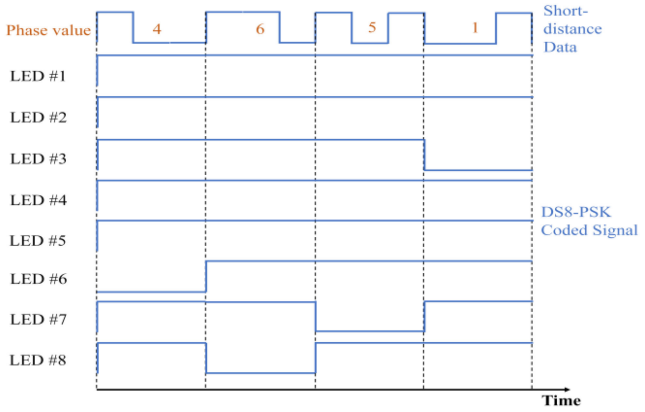

Fig. 3. DS8-PSK signal of an LED data group.

TABLE I
BIT CLUSTER MAPPING TO PHASE [34]

3 bits	Phase shift value
{0 0 0}	0
{0 0 1}	1
{0 1 0}	2
{0 1 1}	3
{1 0 0}	4
{1 0 1}	5
{1 1 0}	6
{1 1 1}	7

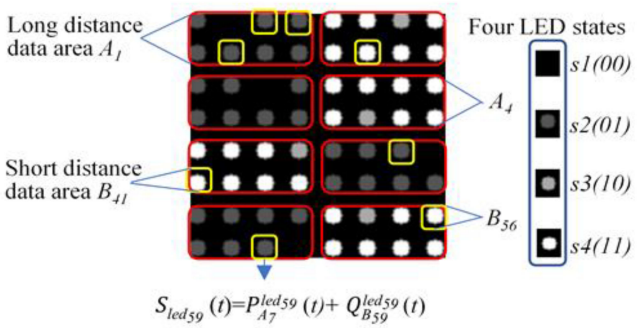

Fig. 4. Hierarchical coding scheme.

relationship under different dimming levels. And there are 7 dimming levels to control the power intensity of the LED matrix. The dimming level *dl* is defined as the ratio of the number of LEDs with the state of '1' to the total number of LEDs in a group of 8 LEDs.

After the short-distance data is coded by DS8-PSK and the long-distance data is Manchester coded, hierarchical coding is used to realize long- and short-distance transmission simultaneously. The hierarchical coding scheme [35] is shown in Fig. 4.

The 8 × 8 LED matrix is divided into different sub-areas to transmit data for long- and short-distance transmission

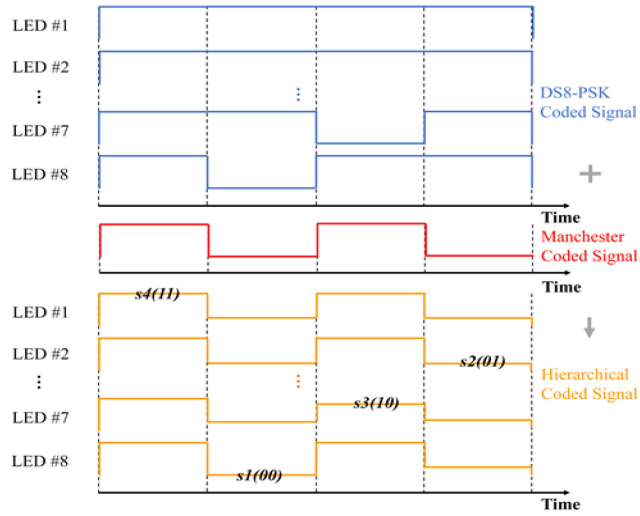

Fig. 5. Hierarchical coded signal of an LED data group.

respectively. The red represents the long-distance data area A_n ($n = 1, 2 \ldots, 8$), and it is the Manchester coded data for long-distance transmission. The yellow represents the short-distance data area B_m ($m = 1, 2 \ldots, 64$), and it is the DS8-PSK coded data for short-distance transmission. In the paper, for 8 × 8 LED matrix, there are 8 data areas for long-distance, and 64 data areas for short-distance. In addition, each long-distance data area of A_n is composed of 2 × 4 LEDs, and it transmits one bit of the Manchester coded data at a time slot. Each short-distance data area of B_m is composed of one LED, and it transmits one bit of DS8-PSK coded data at a time slot. For a LED led_i ($i = 1, 2 \ldots, 64$), the data for long-distance and short-distance transmission at time slot of t are $P_{A_n}^{led_i}(t)$ and$Q_{B_m}^{led_i}(t)$, respectively. Then the output signal S_{ledi} of led_i is obtained by combining $P_{A_n}^{led_i}(t)$ and$Q_{B_m}^{led_i}(t)$, and it is given by:

$$S_{led_i}(t) = P_{A_n}^{led_i}(t) + Q_{B_m}^{led_i}(t) \tag{1}$$

Fig. 5 shows the hierarchical coded signal that overlaps the DS8-PSK coded short-distance data and the Manchester coded long-distance data with 8 LEDs as LED #1 …LED #8. For the Manchester coded data and DS8-PSK coded data, it has two values of '0' and '1'. After combining with them, it is obtained as $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Then it is mapped to four different states as *s1*, *s2*, *s3* and *s4* of the LED, respectively. And it is corresponding to four different gray levels, as shown in Fig. 4. The four gray levels of the LED can be achieved by changing the pulse width by PWM modulation. The transmission power P_t of the LED matrix is expressed as:

$$P_t = \frac{Num}{2} \times [dl \times (P_{s2} + P_{s4}) + (1 - dl) \times (P_{s1} + P_{s3})] \tag{2}$$

where *Num* is the number of LEDs, *dl* is the dimming level, P_{s1}, P_{s2}, P_{s3} and P_{s4} are the power of the four states of the LED,

TABLE II
LED DATA GROUP STATE UNDER DIFFERENT DIMMING LEVEL[34]

Dimming level 1/8	Dimming level 2/8	Dimming level 3/8	Dimming level 4/8	Dimming level 5/8	Dimming level 6/8	Dimming level 7/8	Phase difference
10000000	10000001	10000011	10000111	10001111	10111111	10111111	0
01000000	11000000	11000001	11000011	11000111	11001111	11011111	1
00100000	01100000	11100000	11100001	11100011	11100111	11101111	2
00010000	00110000	01110000	11110000	11110001	11110011	11110111	3
00001000	00011000	00111000	01111000	11111000	11111001	11111011	4
00000100	00001100	00011100	00111100	01111100	11111100	11111101	5
00000010	00000110	00001110	00011110	00111110	01111110	11111110	6
00000001	00000011	00000111	00001111	00011111	00111111	01111111	7

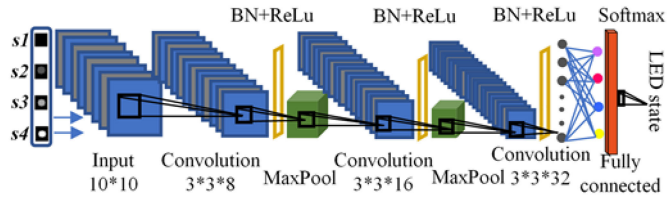

Fig. 6. CNN architecture for LED state classification.

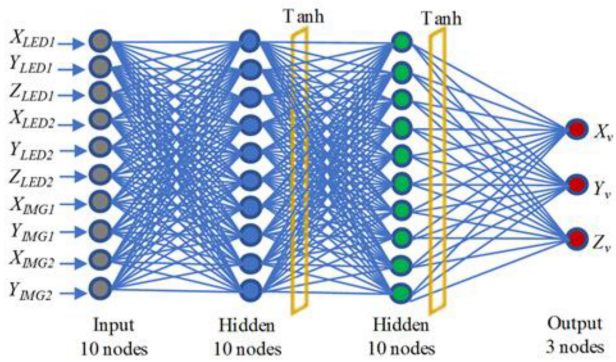

Fig. 7. ANN architecture for vehicle position prediction.

respectively, and it is determined by the power of the LED and the pulse width of the PWM.

B. Receiver

At the receiver, after free space transmission, the captured signal by the IS is presented in the form of an image. After image morphology operation and contour detection, an image with complete LED matrix is obtained. Then it is divided into 64 sub-LED areas, of which 64 LEDs have 4 different states. Subsequently, a four-layer CNN architecture is adopted, as shown in Fig. 6. It includes an input layer, three hidden layers, and an output layer. The data of the input layer is the LED sub-images captured by the camera at different distances, the input size is 10×10, and the number of color channels is 3. Both the first and second hidden layers include a convolutional layer, a batch normalization layer, a ReLu activation layer, and a maximum pooling layer. The third hidden layer contains a convolutional layer, a batch normalization layer, a ReLu activation layer, and a fully connected layer. For the three convolutional layers, the size of filters is 3×3. And the number of filters is 8, 16 and 32, respectively. There are 4 nodes in the output of the fully connected layer, and it is corresponding to the 4 states of the LED. Finally, the classification result of each LED is output through a softmax layer.

After the CNN training is completed, the state of each LED can be obtained, and then hierarchical decoding, Manchester decoding, and DS8-PSK decoding are performed. Finally, the data for long-distance and short-distance transmission is recovered. In the hierarchical decoding, firstly, the state of the LED is demapped, and the DS8-PSK coded short-distance data and Manchester coded long-distance data are obtained respectively. In the paper, since there is one LED in the short-distance data area B_m, and it transmits one DS8-PSK coded data bit, thus, the data can be easily obtained. For the long-distance data area A_n, there are 2×4 LEDs, and it transmits one Manchester coded data bit. Thus, the data $\tilde{P}_{A_n}$ is recovered by judging the average value $\bar{P}_{A_n}^{led_i}$ ($i = 1, 2 \ldots, 8$, $n = 1, 2 \ldots, 8$) of the data in a group of 2×4 LEDs, and it is given by:

$$\tilde{P}_{A_n} = \begin{cases} 1, & if \bar{P}_{A_n}^{led_i} \geq 0.5 \\ 0, & else \end{cases} \tag{3}$$

In the traditional positioning algorithms, it is based on the mapping of 3D coordinates to 2D coordinates according to the visual imaging principle of IS, and large number of LED reference points are required. Thus, its computational cost is high. In addition, it is difficult to achieve 3D coordinate positioning based on similar triangle positioning algorithms. To realize the positioning of the vehicle effectively, in the paper, after obtaining the location data of the two LED matrixes, a three-layer ANN architecture is utilized to predict the position of the vehicle. It is shown in Fig. 7. It consists of an input layer, two hidden layers and an output layer. There are 10 nodes in the input layer, in which are the location data of the two LED matrixes (X_{LED1}, Y_{LED1}, Z_{LED1}), (X_{LED2}, Y_{LED2}, Z_{LED2}) and the location coordinates of the two LED matrixes in the image (X_{IMG1}, Y_{IMG1}), (X_{IMG2}, Y_{IMG2}) is calculated by LED contour and centroid extraction. Moreover, the number of neurons in the two hidden layers is 10, the activation function is Tanh, and there are 3 nodes in the output layer, in which are the world coordinates of the vehicle (X_v, Y_v, Z_v).

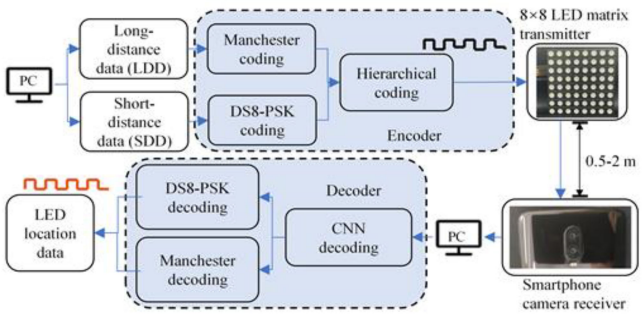

Fig. 8. The experimental setup.

TABLE III
THE KEY EXPERIMENTAL PARAMETERS

LED matrix	Keyes2388ARGB
LED matrix size	6×6 cm
LED number	64
LED diameter	5 mm
LED spacing	3 mm
Image sensor	Sony IMX519
Resolution	1280×720
Frame rate	30 fps
ISO	100
Shutter speed	1/120

III. EXPERIMENTAL SETUP AND RESULT DISCUSSIONS

The experimental setup is shown in Fig. 8. Firstly, according to the geographic distribution of long-distance data area and short-distance data area on the LED matrix in Fig. 4, the LED location data is mapped into long-distance data (LDD) and short-distance data (SDD). Then hierarchical coding based on LDD and SDD is performed, and it is sent through an 8 × 8 LED matrix by PWM. After transmission over free space, the modulated signal is received by the smartphone camera. Subsequently, the LED location data is recovered through the decoder in the PC. In the paper, the LED matrix is used as Keyes2388ARGB, and the size is 6 cm × 6 cm. The number of LEDs is 8 × 8, the diameter of a single LED is 5 mm, and the distance between the LEDs is 3 mm. The IS is used as Sony IMX519, the resolution is 1280 × 720, the frame rate is 30 fps, the ISO is 100, and the shutter speed is 1/120. The experimental parameters are shown in Table III.

To investigate bit error rate (BER) performance of the proposed IS-based VLP scheme, the BER is measured at different distances between the IS and the LED matrix. Fig. 9 shows the BER performance of the proposed IS-based VLP scheme at different distances under three dimming levels of 1/8, 4/8 and 7/8, respectively. In Fig. 9, the BER of the LDD and SDD using the proposed CNN-based decoding scheme, the K-means-based decoding scheme, and the conventional decoding (CD) scheme [35] are given, respectively. For SDD at the distance of 0.5m, as the dimming level is 1/8, 4/8 and 7/8, the BER of the proposed CNN-based decoding scheme is 4.7×10^{-3}, 4.2×10^{-3} and 8.1×10^{-4}, and the BER of K-means-based decoding scheme is 2.8×10^{-1}, 4.9×10^{-1}, and 5.3×10^{-1}, respectively. Based on CD scheme, the BER is 5.8×10^{-1}, 3.9×10^{-1}, and 5.3×10^{-1}, respectively. It shows that the BER performance of the proposed CNN-based decoding scheme under the three dimming levels is better than that of the K-means-based and CD schemes. The reason is that the LED matrix is small-sized in the experiment, and each LED transmits one SSD bit, thus, the number of pixels between each LED in the captured image is very small, and it makes each LED susceptible to the influence of nearby LED light sources or noise. Based on K-means and CD schemes, the fidelity requirements of the LED are high, and it affects the BER performance. In addition, by using CNN, it can learn the characteristics of each LED effectively. As the data trains set is enough, it is robust to the distorted LED images. Therefore, it can achieve better BER performance.

For LDD, the BER performance is affected by two factors as communication distance and SDD interference. As the distance between the IS and the LED matrix is 0.5 m and 1 m under the dimming level of 4/8, the spatial separability of the received signal is strong, and the high-frequency component is enough, thus, the BER of 1.25×10^{-4} can be achieved using CNN-based and CD schemes. As the distance between the IS and the LED matrix is 1.5 m, and the SDD component is distorted, the BER performance of LDD drops sharply. When the distance between the IS and the LED matrix is 2 m, the SDD component is almost lost. As the bit of LDD occupies more pixels on the image, the BER performance of LDD increases as 1.25×10^{-4}. In Fig. 9, when the dimming levels are 1/8, 4/8, and 7/8, and the distance between the IS and the LED matrix is 1.5 m, for the proposed CNN-based decoding scheme, the BER is 1.8×10^{-2}, 2.5×10^{-4}, and 8.8×10^{-4}, respectively. For the K-means-based decoding scheme, the BER is 1.3×10^{-1}, 1.2×10^{-1}, and 1.2×10^{-1}. And for the CD scheme, the BER is 1.25×10^{-4}, 7.9×10^{-3}, and 4×10^{-2}, respectively. Therefore, for LDD using the CNN-based decoding scheme, the BER performance is generally better than that with the K-means-based and CD schemes.

Moreover, from Fig. 9, it can be seen that under different dimming levels and distances, the BER of the LDD using the three decoding schemes is better than the BER of the SDD. This is because each LED transmits one bit of short-distance data, and a group of 8 LEDs together transmit a long-distance data bit. Long-distance data has better spatial high-frequency characteristics, and it is more robust to interference from noise or adjacent LEDs.

Fig. 10 shows the vehicle positioning errors in the three directions of X, Y, and Z axes by using the proposed IS-based VLP scheme with ANN through simulation. The mean positioning error is defined as follows:

$$\begin{bmatrix} X_{Mean\ positioning\ error} \\ Y_{Mean\ positioning\ error} \\ Z_{Mean\ positioning\ error} \end{bmatrix} = \frac{1}{N} \times \sum_{i=1}^{N} \left(\begin{bmatrix} abs(X^i_{ture} - X^i_{pred}) \\ abs(Y^i_{ture} - Y^i_{pred}) \\ abs(Z^i_{ture} - Z^i_{pred}) \end{bmatrix} \right) \quad (4)$$

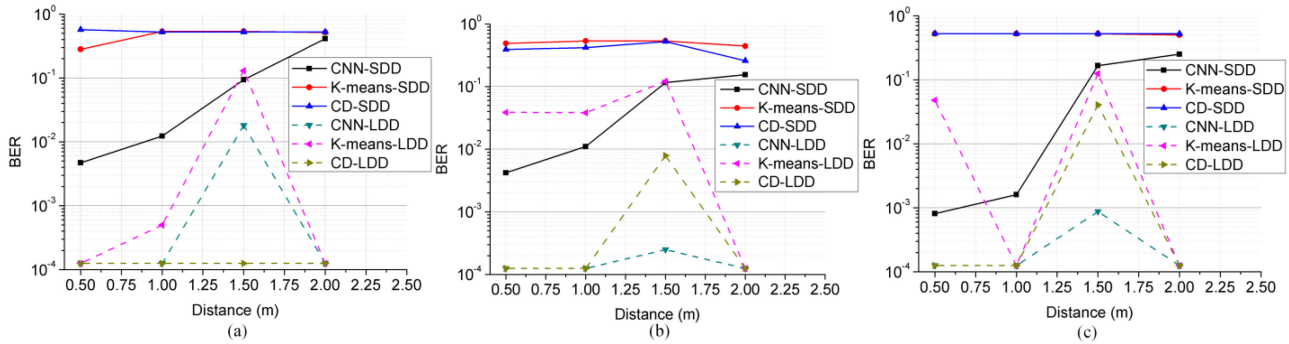

Fig. 9. BER performance at different distances. (a) 1/8 dimming level. (b) 4/8 dimming level. (c) 7/8 dimming level.

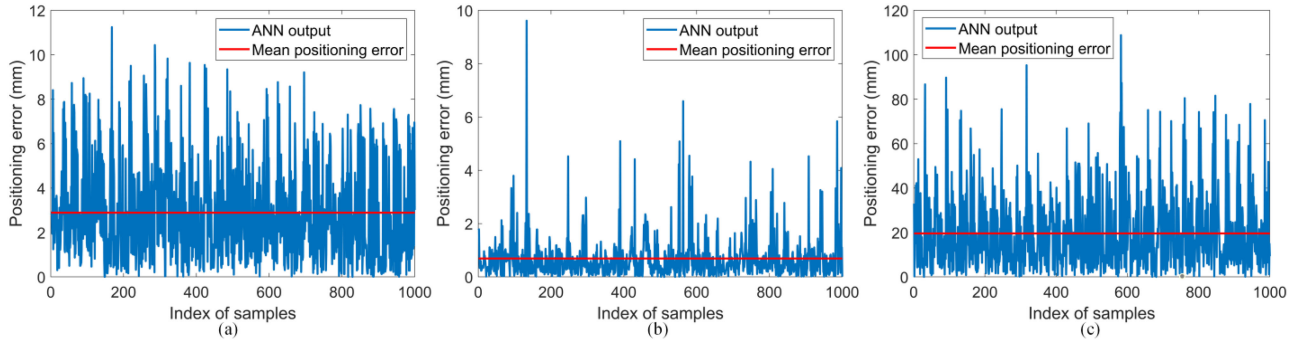

Fig. 10. Vehicle positioning errors using ANN. (a) *X* axis, (b) *Y* axis, (c) *Z* axis.

For *i*-th sample, the positioning error is defined as the absolute value difference between the location coordinates predicted by ANN and the real location coordinates, and the mean positioning error is defined as the sum of positioning error of all samples divided by the number of samples *N*. In the simulation, the *Z*-axis direction is parallel to the vehicle's forward direction, the *Y*-axis is perpendicular to the road plane, and the *X*-axis is perpendicular to the vehicle's forward direction. The two LED matrixes are 6 m apart, and their height from the road surface is 5 m. For the image sensor, it has no tilt or rotation during the movement. The coordinates of the two LED matrixes in the image are obtained according to the principle of geometric imaging. From Fig. 10, it can be seen that for 1000 samples, the average positioning errors of the *X* axis, *Y* axis and *Z* axis are 2.9 mm, 0.7 mm, and 19.6 mm, respectively. And the unbiased estimated standard deviation of the X axis, Y axis and Z axis are 2.2 mm, 0.8 mm, and 16.6 mm, respectively. The total average positioning error is 19.8 mm.

Fig. 11 shows the cumulative distribution function (CDF) of the positioning errors for *X*, *Y*, and *Z* axis. It can be seen that the maximum positioning errors of *X*, *Y*, and *Z* axis are 11.2 mm, 9.6 mm, and 108.8 mm, respectively. The reason for the large *Z*-axis positioning error is that only the use of coordinate information of two LED arrays and one image sensor cannot obtain depth information which is the value of the *Z*-axis in this simulation. By using ANN, it can only learn part of the characteristics of this nonlinear mapping through training process. Therefore, relative

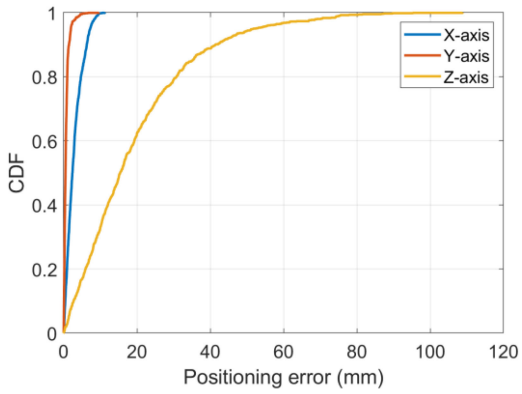

Fig. 11. CDF of positioning errors.

to the *X* and *Y* axes, the *Z* axis coordinate output by ANN has a larger positioning error.

IV. Conclusion

In the paper, the neural network-assisted vehicle positioning scheme based on dimmable LED traffic lights and IS was proposed and demonstrated. Due to the performance of the communication and positioning using VLC based on IS deteriorated rapidly in the long-distance transmission, at the transmitter, the hierarchical encoding method was proposed to support long- and

short-distance data transmission at the same time. In addition, it could also support the dimming of smart LED traffic lights. At the receiver, a CNN was used to classify the state of each LED to decode the received signal, and an ANN was used to predict the position of the vehicle. The results showed that as the transmission distance was 2 m, the BER performance using the proposed scheme was 1.25×10^{-4}. And the average positioning error was 19.8 mm at the maximum simulated distance of 30 m.

REFERENCES

[1] Z. Chen and X. Huang, "Pedestrian detection for autonomous vehicle using multi-spectral cameras," *IEEE Trans. Intell. Veh.*, vol. 4, no. 2, pp. 211–219, Jun. 2019.

[2] C. H. Nguyen, V. H. Nguyen, and Y. M. Jang, "Optical camera communication (OCC) applications for internet of vehicle (IoV)," in *Proc. Int. Conf. Inf. Commun. Technol. Convergence*, 2019, pp. 512–514.

[3] Q. Liang, J. Lin, and M. Liu, "Towards robust visible light positioning under LED shortage by visual-inertial fusion," in *Proc. Int. Conf. Indoor Positioning Indoor Navigation*, 2019, pp. 1–8.

[4] A. Căilean and M. Dimian, "Current challenges for visible light communications usage in vehicle applications: A survey," *IEEE Commun. Surv. Tut.*, vol. 19, no. 4, pp. 2681–2703, Oct.–Dec. 2017.

[5] J. He *et al.*, "Efficient sampling scheme based on length estimation for optical camera communication," *IEEE Photon. Technol. Lett.*, vol. 31, no. 11, pp. 841–844, Jun. 2019.

[6] J. He, Z. Jiang, J. Shi, Y. Zhou, and J. He, "A novel column matrix selection scheme for VLC system with mobile phone camera," *IEEE Photon. Technol. Lett.*, vol. 31, no. 2, pp. 149–152, Jan. 2019.

[7] J. He, Y. Yang, and J. He, "Artificial neural network-based scheme for 4-PWM OCC system," *IEEE Photon. Technol. Lett.*, vol. 34, no. 6, pp. 333–336, Mar. 2022.

[8] Z. Huang, J. He, K. Yu, and W. Li, "Efficient demodulation scheme based on adaptive clock extraction and mapping-sampling for a mobile OCC system," *Appl. Opt.*, vol. 60, no. 12, pp. 3308–3313, 2021.

[9] J. Shi, J. He, J. He, Z. Jiang, Y. Zhou, and Y. Xiao, "Enabling user mobility for optical camera communication using mobile phone," *Opt. Exp.*, vol. 26, no. 17, pp. 21762–21767, 2018.

[10] J. He, K. Yu, Z. Huang, and Z. Chen, "Multi-column matrices selection combined with k-means scheme for mobile OCC system with multi-LEDs," *IEEE Photon. Technol. Lett.*, vol. 33, no. 12, pp. 623–626, Jun. 2021.

[11] K. Yu, J. He, and Z. Huang, "Decoding scheme based on CNN for mobile optical camera communication," *Appl. Opt.*, vol. 59, no. 23, pp. 7109–7113, 2020.

[12] L. Liu, R. Deng, J. Shi, J. He, and L. Chen, "Beyond 100-Kbit/s transmission over rolling shutter camera-based VLC enabled by color and spatial multiplexing," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, 2020, pp. 1–3.

[13] P. Luo, Z. Ghassemlooy, H. L. Minh, X. Tang, and H. M. Tsai, "Undersampled phase shift ON-OFF keying for camera communication," in *Proc. 6th Int. Conf. Wireless Commun. Signal Process.*, Hefei, China, 2014, pp. 1–6.

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

[15] R. D. Roberts, "A MIMO protocol for camera communications (CamCom) using undersampled frequency shift ON-OFF keying (UFSOOK)," in *Proc. IEEE Globecom Workshops*, Atlanta, GA, USA, 2013, pp. 1052–1057.

[16] H. Aoyama and M. Oshima, "Visible light communication using a conventional image sensor," in *Proc. 12th Annu. IEEE Consum. Commun. Netw. Conf.*, Las Vegas, USA, 2015, pp. 103–108.

[17] S. H. Chen and C. W. Chow, "Color-shift keying and code-division multiple-access transmission for RGB-LED visible light communications using mobile phone camera," *IEEE Photon. J.*, vol. 6, no. 6, Dec. 2014, Art. no. 7904106.

[18] Y. Zhuang *et al.*, "A survey of positioning systems using visible LED lights," *IEEE Commun. Surv. Tut.*, vol. 20, no. 3, pp. 1963–1988, Jul.–Sep. 2018.

[19] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, "A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 829–846, Apr. 2018.

[20] C. Ang, "Vehicle positioning using WIFI fingerprinting in urban environment," in *Proc. IEEE 4th World Forum Internet Things*, 2018, pp. 652–657.

[21] J. Luo, L. Fan, and H. Li, "Indoor positioning systems based on visible light communication: State of the art," *IEEE Commun. Surv. Tut.*, vol. 19, no. 4, pp. 2871–2893, Oct.–Nov. 2017.

[22] X. Liu, X. Wei, and L. Guo, "DIMLOC: Enabling high-precision visible light localization under dimmable LEDs in smart buildings," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3912–3924, Apr. 2019.

[23] Y. Li, Z. Ghassemlooy, X. Tang, B. Lin, and Y. Zhang, "A VLC smartphone camera based indoor positioning system," *IEEE Photon. Technol. Lett.*, vol. 30, no. 13, pp. 1171–1174, Jul. 2018.

[24] J. Lain, L. Chen, and S. Lin, "Indoor localization using K-pairwise light emitting diode image-sensor-based visible light positioning," *IEEE Photon. J.*, vol. 10, no. 6, Dec. 2018, Art. no. 7909009.

[25] B. W. Kim and S. Jung, "Vehicle positioning scheme using V2V and V2I visible light communications," in *Proc. IEEE 83rd Veh. Technol. Conf.*, 2016, pp. 1–5.

[26] M. S. Ifthekhar, N. Saha, and Y. M. Jang, "Stereo-vision-based cooperative-vehicle positioning using OCC and neural networks," *Opt. Commun.*, vol. 352, pp. 166–180, 2015.

[27] M. T. Hossan *et al.*, "A new vehicle localization scheme based on combined optical camera communication and photogrammetry," *Mobile Inf. Syst.*, vol. 2018, pp. 1–14, 2018.

[28] J. He and B. Zhou, "Vehicle positioning scheme based on visible light communication using a CMOS camera," *Opt. Exp.*, vol. 29, no. 17, pp. 27278–27290, 2021.

[29] T. H. Do and M. Yoo, "Visible light communication based vehicle positioning using LED street light and rolling shutter CMOS sensors," *Opt. Commun.*, vol. 407, pp. 112–126, 2018.

[30] T. Do and M. Yoo, "Visible light communication-based vehicle-to-vehicle tracking using CMOS camera," *IEEE Access*, vol. 7, pp. 7218–7227, 2019.

[31] X. Zhao and J. M. Lin, "Maximum likelihood estimation of vehicle position for outdoor image sensor-based visible light positioning system," *Opt. Eng.*, vol. 55, no. 4, 2016, Art. no. 043104.

[32] V. T. B. Tram and M. Yoo, "Vehicle-to-vehicle distance estimation using a low-resolution camera based on visible light communications," *IEEE Access*, vol. 6, pp. 4521–4527, 2018.

[33] J. He, K. Tang, J. He, and J. Shi, "Effective vehicle-to-vehicle positioning method using monocular camera based on VLC," *Opt. Exp.*, vol. 28, no. 4, pp. 4433–4443, 2020.

[34] T. Pham, H. Nguyen, T. Nguyen, and Y. M. Jang, "A novel neural network-based method for decoding and detecting of the DS8-PSK scheme in an OCC system," *Appl. Sci.*, vol. 9, no. 11, 2019, Art. no. 2242.

[35] S. Nishimoto, T. Yamazato, H. Okada, T. Fujii, T. Yendo, and S. Arai, "High-speed transmission of overlay coding for road-to-vehicle visible light communication using LED array and high-speed camera," in *Proc. IEEE Globecom Workshops*, 2012, pp. 1234–1238.