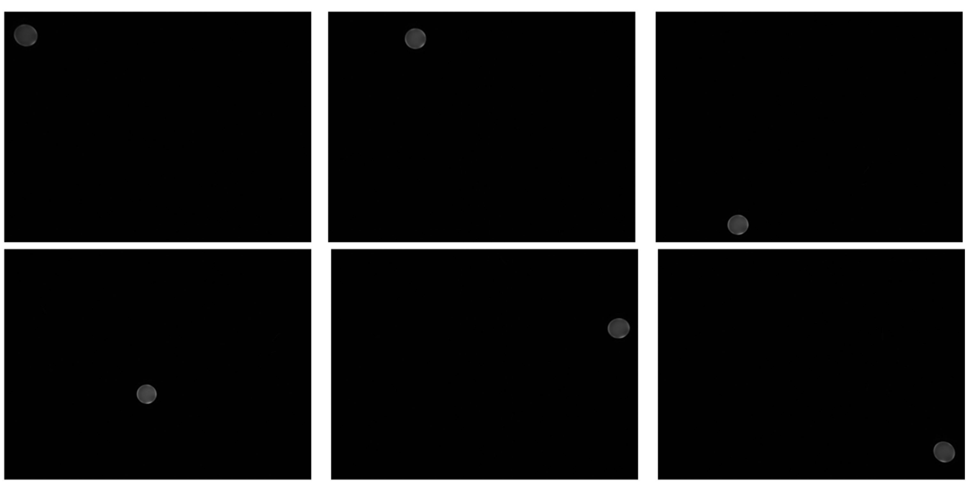
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

A Tilt Receiver Correction Method for Visible Light Positioning Using Machine Learning Method

Volume 10, Number 6, December 2018

Tao Yuan
Yiqin Xu
Yong Wang
Peng Han
Junfang Chen



DOI: 10.1109/JPHOT.2018.2880872
1943-0655 © 2018 IEEE

A Tilt Receiver Correction Method for Visible Light Positioning Using Machine Learning Method

Tao Yuan,[1] **Yiqin Xu**,[1,2] **Yong Wang**,[1] **Peng Han**,[1] **and Junfang Chen**[1]

[1]South China Normal University, Guangzhou 510640, China
[2]Guangdong Institute of Semiconductor Industrial Technology, Guangzhou 510650, China

DOI:10.1109/JPHOT.2018.2880872

Manuscript received October 19, 2018; accepted November 7, 2018. Date of publication November 12, 2018; date of current version November 29, 2018. This work is supported by the National Natural Science Foundation of China under Grant 61371176, and the Science and Technology Project of Guangzhou China under Grant 201704020137 and 201707010485. *(Tao Yuan and Yiqin Xu contributed equally to this work.)* Corresponding authors: Peng Han and Junfang Chen (e-mail: hanp@scnu.edu.cn; gdsled@126.com).

Abstract: In recent years, a visible light positioning (VLP) technology based on complementary metal-oxide-semiconductor sensors has been widely studied due to its high precision and high robustness. However, the existing VLP algorithm based on image sensors often fails to achieve a good positioning effect when the camera is tilted. In order to solve this problem, we propose a neural network algorithm to correct the error caused by the tilt angle of the camera. Because when the tilt angle is different, the LED image captured by the camera will be different and produce different characteristics. By extracting these features and using neural networks to establish the relationship between the characteristics of the LED image and the distance between the receiving and sending terminal, we finally achieve the positioning of the camera by the triangulation algorithm. Experiments demonstrate that our positioning algorithm can achieve high-precision positioning and can be applied to most indoor positioning systems.

Index Terms: Neural network algorithm, tilt angle, visible light positioning.

1. Introduction

Nowadays, with the continuous development of economy and the advancement of contemporary technology, the demand for indoor navigation and positioning is increasing, such as indoor service robots, indoor parking, indoor location-based services (LBS). Traditional indoor positioning methods, such as WLAN, ZigBee, Bluetooth, infrared positioning, ultrasonic positioning, the positioning error reaches tens of centimeters or even meters, and are also susceptible to electromagnetic interference [1]–[3]. Therefore, an indoor positioning technology based on visible light communication has been proposed.

Visible light positioning technology can be divided into two modes, one is non-imaging receiving positioning, and the other is imaging receiving positioning. Non-imaging receiving positioning refers to using a Photodiode (PD) as the receiving terminal of a visible light signal, and positioning is achieved by receiving the light intensity of different LEDs through PD to calculate the distance between PD and different LEDs. This method cost less, but it is easily affected by environmental background light and the reflection effect of objects such as interior furniture, walls, which makes the received signal appear distorted, and reduce its robustness. Imaging reception positioning refers to the use of the camera's image sensor as the receiving terminal of the positioning system. The indoor positioning is performed by the image processing technology and the geometric relationship between the receiving image and the LED. This method is nearly not affected by the ambient light and the background light, and its positioning accuracy is up to centimeters. People have done a lot of work in imaging receiving and positioning. In [4], Jae-Yoon Kim *et al.*, adopts the method of dual-lamp positioning to achieve three-dimensional positioning with an error of less than 6.5 cm. In [5], a geometry-based three-lamp positioning algorithm was proposed, with a positioning accuracy error of up to 1 cm. In [6], high-precision two-dimensional and three-dimensional positioning was achieved by using the same three-lamp positioning algorithm. In Ref. [4]–[6], the real position of the LED and the geometric relationship of the image are used to achieve the positioning. The premise of the positioning algorithm is that the camera must be placed horizontally, that is, the z-axis of the camera coordinate system is parallel to the z-axis of the world coordinate system. But in most practical situations, the z-axis of the camera coordinate system as the receiving terminal cannot always be parallel to the z-axis of the world coordinate system. In [7], a single-lamp positioning algorithm based on geometric relations was proposed. According to the geometric characteristics of the LED, the author derives the mapping relationship between the camera coordinate system and the world coordinate system. It is suitable for the case where the z-axis of the camera coordinate system is not parallel to the z-axis of the world coordinate system, and its positioning accuracy can reach 17.52 cm. In [8], in order to solve the problem of the rotation angle at the receiving end, a VLP algorithm integrating an acceleration sensor with an image sensor was proposed. The rotation angle of the receiving end obtained by the acceleration sensor and the LED image received by the image sensor are used for positioning, and its positioning accuracy can reach 10 cm. In [9], the combination of an image sensor and a motion sensor achieves a positioning accuracy of 4.4 cm. In [7]–[9], although the receiver tilt angle problem has been solved, the positioning error in the literature [7] is relatively large and the calculation method is complicated. And although the positioning accuracy is high in the literature [8], [9], additional sensors are needed, which increases the cost and complexity of the system.

In this paper, a high-precision VLP positioning algorithm based on machine learning is proposed. Different from the traditional algorithm, we obtain the relative distance between the camera of the receiving end and the LED of the transmitting end through the machine learning method, and then obtain the position of the receiving end through the triangulation algorithm. First, a circular LED downlight is horizontally mounted on the roof as the VLP's position signal transmitter. When the distance between the receiver camera and the transmitting LED or the relative receiving angle changes, the received LED image will be deformed into an ellipse, and the characteristics of the ellipse will also change accordingly. So from the image, we can get 5 features, namely, the area of the ellipse image, the major axis of the ellipse, the minor axis of the ellipse, and two position parameters x, y of the ellipse centroid in the pixel coordinate system. In order to use the extracted features to obtain the distance between the camera and the LED, we choose the MLP neural network to achieve. The MLP is not the best performing ANN; However, it is the least complex and supports the Levenberg-Marquardt back-propagation (BP) algorithm, which is very popular due to the ease of hardware implementation [15]. Our experiments show that the positioning algorithm proposed in this paper can achieve positioning accuracy of 1.9 cm in the case of 50 hidden layers of neural network. This method achieves centimeter-level high-precision positioning without increasing other costs, and therefore has a broad application prospect.

2. System Principle

2.1 The Relationship Between the Camera Tilt Angle and the Captured Image

In this paper, the circular LED downlight is placed parallel to the x-y plane of the world coordinate system, and the receiving camera captures the LED image at a constant position. Because the camera does not stay horizontal at all times, once the camera tilts, the captured LED image will be

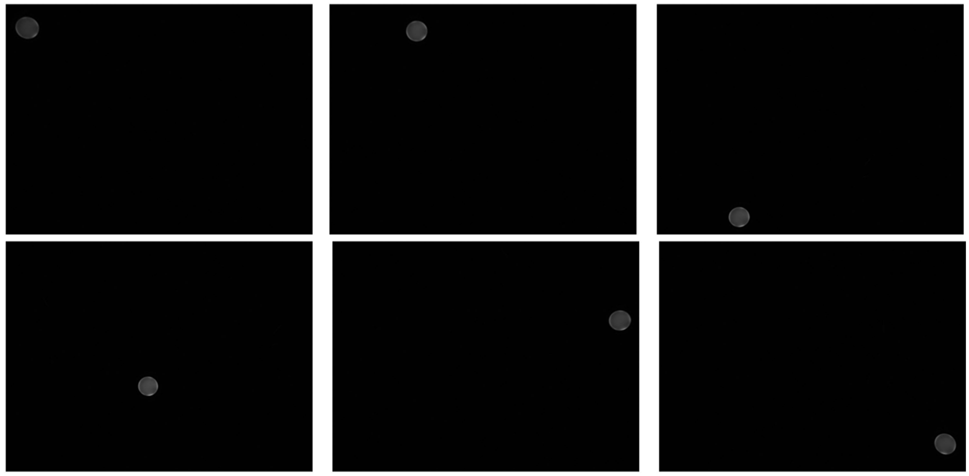


Fig. 1. The LED image captured with different tilt angles.

different accordingly. Fig. 1 shows the images taken with the camera at the same position and with different tilt angles.

From the image we can see that the circular LED downlight is mapped to an oval LED image. When the distance between the camera on the receiving end and the LED is constant, and when the tilt angle of the camera changes, the coordinates of the centroid of the captured elliptical LED image on the pixel coordinate system will change, the major and minor axes of the ellipse will change as well. In addition, according to the principle of small hole imaging, when the distance between the camera and the LED increases, the area of the LED pattern captured by the camera will become smaller. Therefore, we select five features of the LED image: the x coordinate of the centroid of the LED image, the y coordinate of the centroid of the LED image, the length of the major axis of the LED, the length of the minor axis of the LED, and the area of the ellipse LED image, as input features and acquire the distance between the camera and the LED through the machine learning method.

2.2 LED Image Feature Extraction

As we mentioned in 2.1, we need to select the x axis coordinate, the y axis coordinate, the length of the long axis, the length of the minor axis, and the area of the ellipse LED image as input features, and get the distance between camera and LED through machine learning. The steps for feature extraction are as follows, Fig. 2 shows the flow diagram of the feature extraction.

Step 1: LED area selection

To obtain the various features of the LED, the first step is to determine the location of the LED. First of all, we convert the obtained image into a grayscale image and binarize the grayscale image. Then we select the region of interest (ROI) and divide the LED image to obtain the subimage we need, as shown in Fig. 3.

Step 2: Acquisition of LED image coordinates

We use the ellipse fitting algorithm in the literature [10] to fit our LED image and get the ellipse fitting curve of the LED image. The center of mass of the elliptical image of the LED is thus obtained. And from the position of the LED ellipse image determined in Step 1 in the entire image, we can obtain the coordinates (x, y) of the centroid of the LED image in the pixel coordinate system.

Step 3: Acquisition of LED image area

The binarized image has only two values, 0 and 1. From the graph 3, only the LED image has a value of 1, and the other background has a value of 0. Therefore, the area of the LED image can be converted into the number of pixels with a value of 1 in the image after binarization.

Step 4: Acquisition of major and minor axes of LED ellipse images

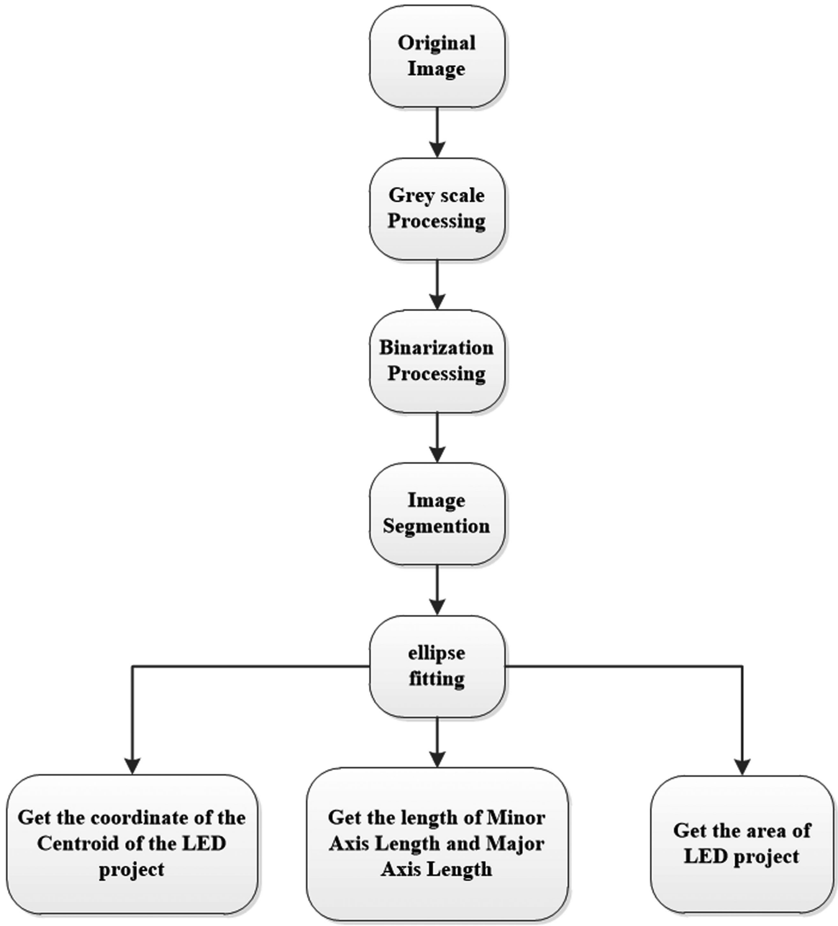


Fig. 2. The flow diagram of the feature extraction.

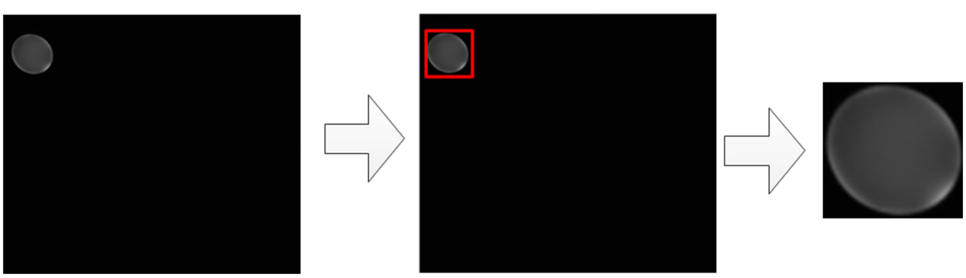


Fig. 3. The schematic of the LED image segmentation.

In Step 2, we have obtained the fitting curve of the LED image. Therefore, the major and minor axes of the ellipse can be obtained from the fitting curve, as shown in Fig. 4.

2.3 Neural Network

With the change of the distance between the camera and the LED and the tilt angle of the camera, the five characteristics of the LED image obtained by taking a picture also change. Therefore, in this paper, we consider establishing the relationship between the characteristics of the LED image and the distance between the receiving end and the sending end. Neural network is a powerful

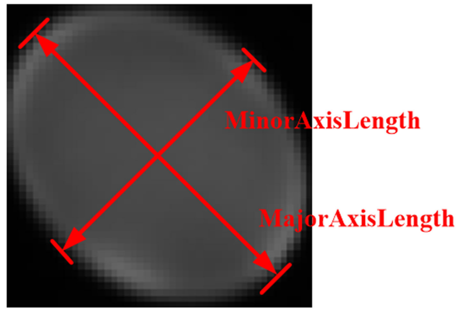


Fig. 4. The major and minor axes of the LED image.

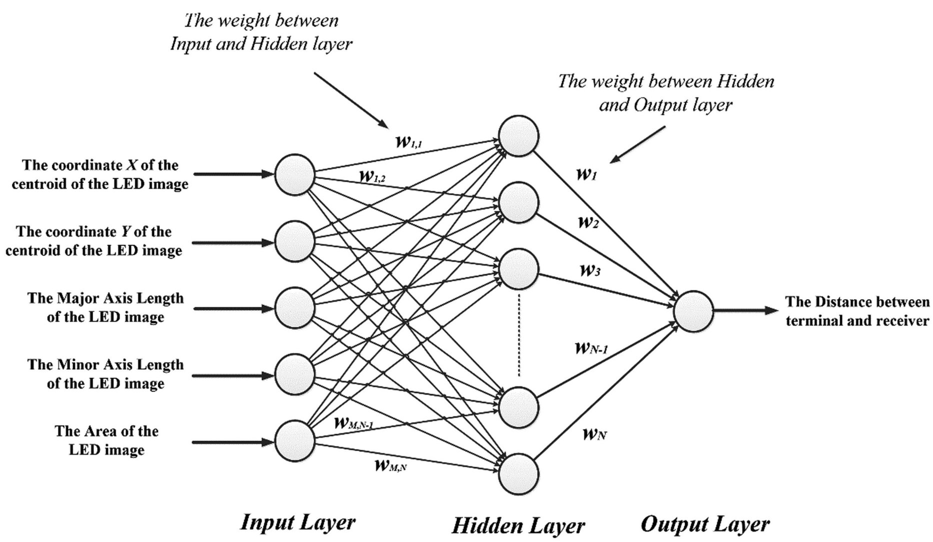


Fig. 5. The structure of the MLP network.

nonlinear mapping tool, and the simplest neural network is Multilayer Perceptron Network (MLP) [11]. Through training, MLP can theoretically realize any nonlinear mapping relationship. Therefore, we use neural networks to establish the relationship between the characteristics of the LED image and the distance between the receiving end and the sending end.

2.3.1 Neural Network Structure: As can be seen from the Fig. 5, the MLP network is composed of a multi-layer structure, including an input layer, an output layer, and a plurality of intermediate layers commonly referred to as hidden layers.

The structure of a single hidden layer node is shown in Fig. 6.

The mathematical description of the relationship between the input and output of each hidden layer node is:

$$z`(k) = f[z(k)] = f\left\{\sum_{j=1}^{N} w_{i,j}(k)\, u_i(k) + b_j\right\} \tag{1}$$

Among them, $w_{i,j}(k)$ is the weight coefficient between the output signal $u_i(k)$ connecting the ith neuron of the input layer and the input signal of the jth neuron of the hidden layer. And $j = 1, 2, \ldots, N$, which means there are a total of N hidden layer nodes. Each constant b_j is the deviation value of the i-th hidden layer, also called the threshold. $f\{\cdot\}$ is the activation function of the hidden layer node. By reasonably selecting weights, thresholds, and nonlinear activation functions, the MLP network can implement any desired non-linear mapping.

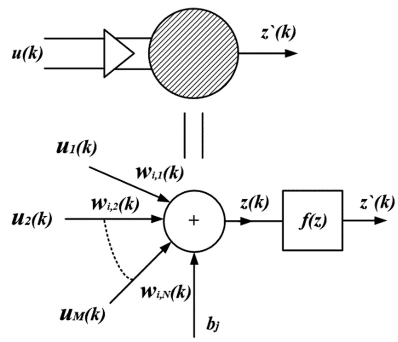


Fig. 6. Node of hidden layer.

The activation function f{·} used in this paper is a Sigmoid function defined as:

$$\mathrm{sgm}\,(x) = \frac{2c_1}{1 + e^{-c_2 x}} - c_1 \tag{2}$$

where c1 and c2 are appropriately selected constants. The derivative of the Sigmoid function is:

$$\mathrm{sgd}\,(x) = \frac{c_1}{2c_1}\left[c_1^2 - \mathrm{sgm}^2\,(x)\right] \tag{3}$$

2.3.2 BP Algorithm: Back propagation (BP) is the most widely used training algorithm for MLP [26]. During training, a set of input values and expected values matching them are given to the neural network, and the link weights are adjusted according to this set of training data. After comparing the expected value with the positively propagated output value, the objective function, namely, the instantaneous output squared error of the neural network, can be obtained:

$$Ek = [yk - d\,(k)]^2 \tag{4}$$

Here, y(k) represents the kth output signal; d(k) is the expected output signal corresponding to y(k). Using the steepest gradient descent method for error backpropagation and weight correction, the error signal can reach or fall below the set value through repeated learning. The correction of the weight is achieved by the following learning equation.

$$\omega_{ij}(n+1) = \omega_{ij}(n) - \gamma \frac{\partial Ex}{\partial \omega_{ij}(n)} \tag{5}$$

Among them, $\omega_{ij}(n)$ represents the connection weighting value, γ represents the learning rate. The choice of learning rate is crucial to the performance of the neural network. Excessive learning rate may cause the convergence of the error to fluctuate, while too small a learning rate leads to a slow convergence rate. With equations (4) and (5), the connection weights can be adjusted to minimize the mean square error between the actual output and the desired output of the neural network. After BP algorithm training, for any input value, the MLP can give a relatively appropriate output.

2.4. Positioning Method

In this system, ordinary round LED lights are installed on the ceiling. Each LED lamp is assigned a unique ID that is associated with its location and stored in the ID location database. The LED periodically transmits its ID information through OOK modulation. The camera uses the Rolling Shutter (RS) to capture the signal sent by the LED and perform demodulation to obtain the position information of the corresponding LED. Because LED is the main source of illumination for indoor lighting, if we set the parameters of the camera appropriately, the projects in the captured image would only be the LEDs, and different LEDs would not affect each other in the image-sensor VLP case [13].

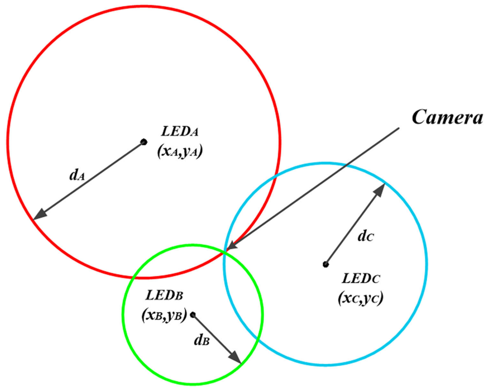


Fig. 7. The principle of the triangulation algorithm.

TABLE 1
Parameter of the Experiment in This Paper

Parameter	Value
The focal length/mm	3
The resolution of the camera	2048×1536
The exposure time of the camera/ms	0.1
The ISO of the camera	100
The pixel size of the camera/μs	3.2×3.2
The diameter of the LED downlight/cm	15
The Power of each LED/W	3
Indoor space unit size (L×W×H)/m	130x130x200

After we get the distance between the camera and the three different LEDs, we can locate the camera with the triangulation algorithm. As shown in the Fig. 7, the coordinates of the three LEDs are known as (x_A, y_A), (x_B, y_B) , and (x_C, y_C) .

Then according to the three circles intersect at one point, solve the equations

$$\begin{cases} (x_e - x_A)^2 + (y_e - y_A)^2 = d_{A,x,y}^2 \\ (x_e - x_B)^2 + (y_e - y_B)^2 = d_{B,x,y}^2 \\ (x_e - x_C)^2 + (y_e - y_C)^2 = d_{C,x,y}^2 \end{cases} \tag{6}$$

The system of equations can be

$$\begin{cases} 2x_e(x_A - x_C) + x_C^2 - x_A^2 + 2y_e(y_A - y_C) + y_C^2 - y_A^2 = d_{C,x,y}^2 - d_{A,x,y}^2 \\ 2x_e(x_B - x_C) + x_C^2 - x_B^2 + 2y_e(y_B - y_C) + y_C^2 - y_B^2 = d_{C,x,y}^2 - d_{B,x,y}^2 \end{cases} \tag{7}$$

After solving, we can get camera coordinates (x_e, y_e) .

3. Experiment Setup

Based on the above theory, we did experiments to verify the feasibility of the system. We built a 130 cm $*$ 130 cm $*$ 200 cm positioning platform. Five LED's are placed on the top of the platform to send different ID information. At the bottom of the platform, the camera is mounted on a bracket with adjustable tilt angle. The camera sends the captured image data to the computer and processed by Matlab software. Finally, we get the positioning result. The experimental setup parameters are shown in Table 1 below. And the whole system of our experiment is shown in Fig. 8.

Fig. 8. The positioning system setup.

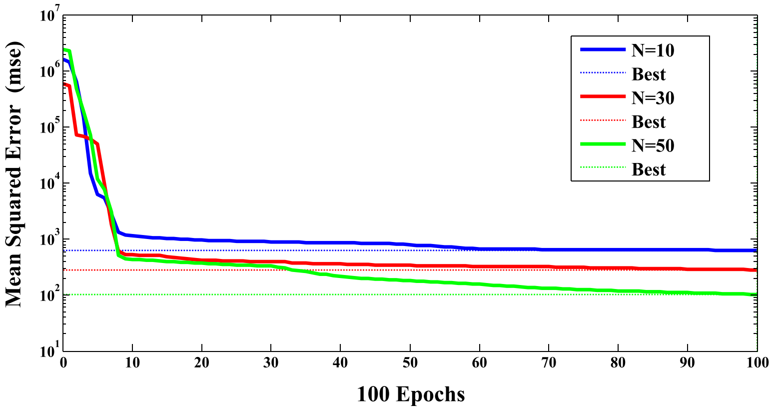


Fig. 9. The training performance with different hidden layer nodes.

3.1 Neural Network Training Error and Hidden Layer Nodes

The number of hidden layer nodes in the neural network affects the performance of the neural network. In order to explore the relationship between neural network hidden layer nodes and neural network training effect, we have selected neural networks with hidden layer nodes $N = \{10, 30, 50\}$ to map the distance relationship between 5 features of the LED and the transmitting receiver. In this experiment, we first conducted the selection of training data. We chose the distance between the transmitter and the receiver from 10 cm to 160 cm, with an interval of 10 cm. Then we collected data for these 16 different distances, collected 100 images at each distance, and processed them to extract the 5 features of the LED image. Finally, these 5 features were sent to the neural network for training. The training results are shown in the Fig. 9.

From the figure we can see that when the number of neural network hidden layer nodes is 10, 30, 50, the neural network training error is 5E3, 2E3, 0.1E2 respectively, which means that with the increase of the neural network hidden layer nodes, the mapping effect of the neural network is better.

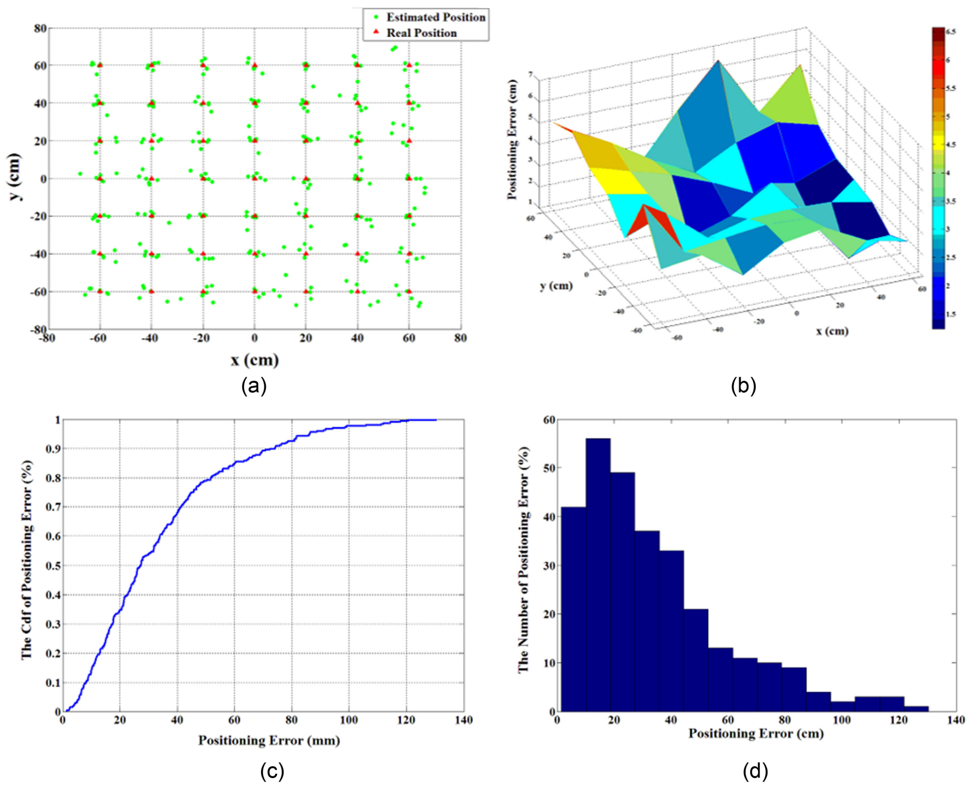


Fig. 10. The positioning results of the system with 10 hidden layer nodes: (a) the positioning result; (b) the location error distribution chart; (c) the CDF curve of the positioning error; (d) the error distribution histograms.

However, the increase of nodes in the hidden layer will inevitably increase the computational load of the neural network, resulting in an increase in the time for each positioning, which will affect the real-time positioning. However, too few hidden layer nodes can affect the accuracy of positioning, and increase the system error. Therefore, while considering the training accuracy of the neural network, we must design the number of hidden layer nodes of the neural network according to the actual situation.

3.2 The Relationship Between Neural Nodes and Positioning Accuracy

After the camera collects the LED photos, it can extract the characteristics of the LED image through the image processing method, and through the trained neural network, we can obtain the distance between the camera and the LED. Once we obtain the distance between the same camera with respect to the three LEDs and determine the coordinates of the LED by decoding the ID of the LED, the coordinate information of the camera can be estimated by the triangulation algorithm. We tested the positioning effect obtained by the proposed positioning algorithm when the number of hidden layer neurons was 10, 30, and 50 respectively. The test results are as follows:

1) *Number of hidden layer nodes = 10:* Fig. 10(a) is the relationship between the spatial location point and the real point measured when the number of nodes in the hidden layer is 10, where the red point is the real location point and the green point is the location estimation point.

In order to avoid random errors, we tested 6 times at each positioning point. The resulting positioning effect is shown in Figure(a). Comparing the estimated point with the real point, the positioning error is obtained, and the location error distribution chart is shown in figure(b). In order to more intuitively demonstrate our positioning effect, we have plotted the error cumulative

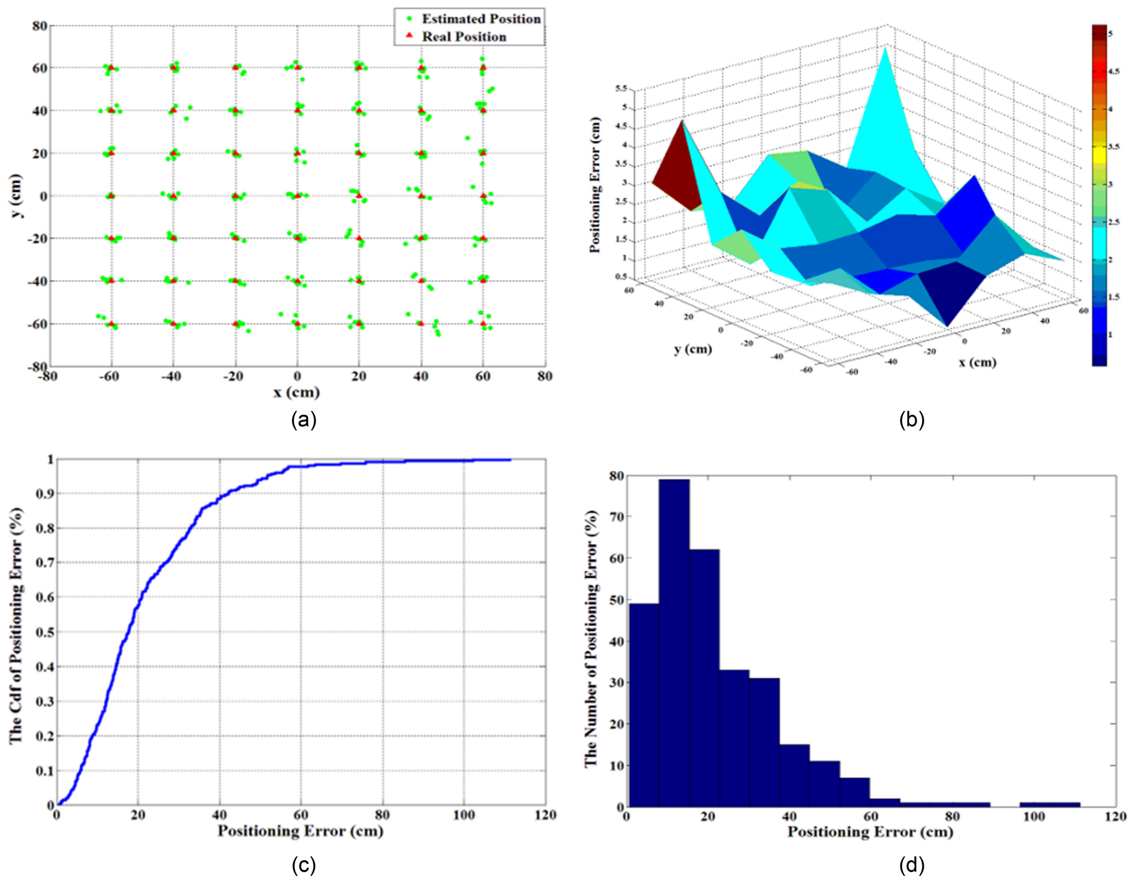


Fig. 11. The positioning results of the system with 30 hidden layer nodes: (a) the positioning result; (b) the location error distribution chart; (c) the CDF curve of the positioning error; (d) the error distribution histograms.

distribution function (CDF) curve and distribution histograms as shown in (c)(d). We can see that more than 90% of the point positioning error is within 7.5 cm.

2) *Number of hidden layer nodes* = *30:* In the same way as (1), where the number of nodes in the hidden layer is 30, we have plotted the orientation, the error space distribution, the error cdf, and the histogram of the error distribution in Fig. 11. And we can see that more than 90% of the point positioning error is within 4.2 cm.

3) *Number of hidden layer nodes* = *50:* In the case where the hidden layer node is 50, we draw the schematic of the location, the error space distribution, the error cdf, and the histogram of the error distribution Fig. 12, we can see that 90% of the point positioning error is within 3.5 cm.

From the above experimental results, we can get that using the neural network for visible light visual positioning can correct the positioning error caused by the tilt of the camera and obtain centimeter-level positioning accuracy. When the number of neurons in the hidden layer neural network increases, the training error of the neural network decreases, and the positioning accuracy also improves. Taking the mean value of each point's positioning error, we can obtain that when the hidden layer nodes are 10, 30, and 50, the average positioning error is 3.6 cm, 2.1 cm, and 1.9 cm, respectively. This positioning accuracy is suitable for most occasions at present, and because there is no need to attach extra sensors, only the image sensor is needed, so the cost of this positioning method is not particularly high.

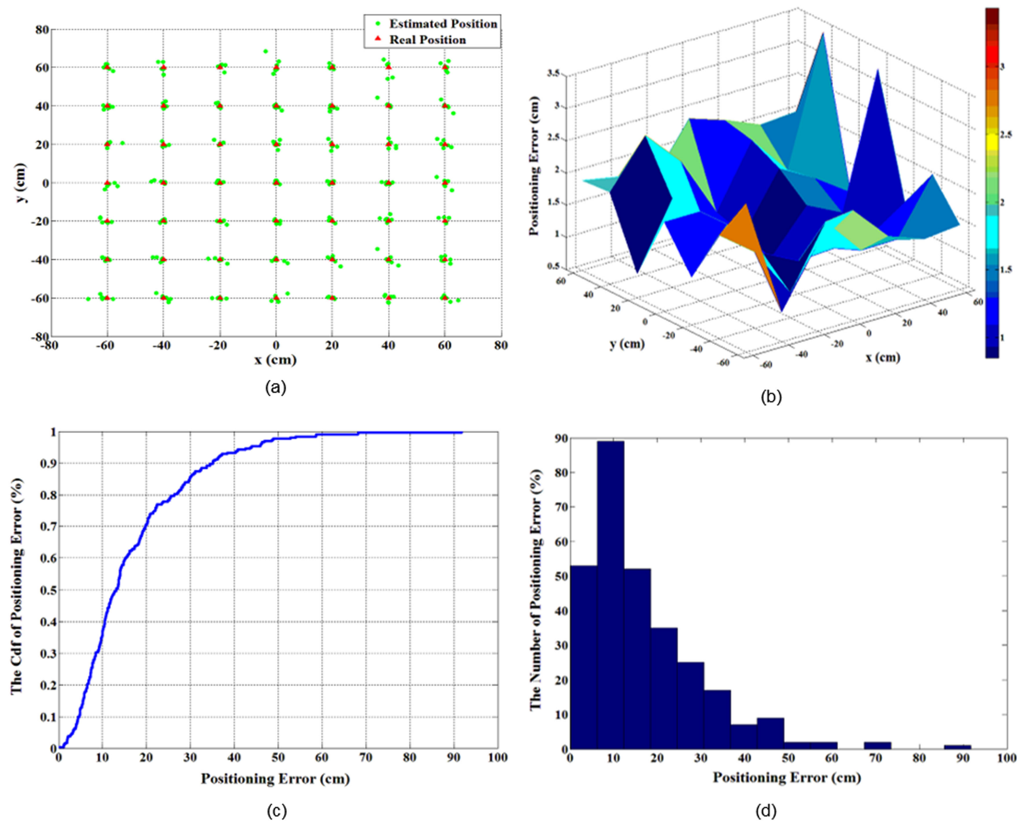


Fig. 12. The positioning results of the system with 50 hidden layer nodes: (a) the positioning result; (b) the location error distribution chart; (c) the CDF curve of the positioning error; (d) the error distribution histograms.

4. Conclusion

In this paper, we propose a positioning method based on machine learning to solve the problem of different received LED images caused by tilt when the camera is used as a visible light receiving terminal. After extracting five features of its image (x-axis coordinates, y-axis coordinates, length of the major axis, minor axis length, area of the elliptical LED image), and obtaining the relative distance between the camera and the LED by mapping from the neural network, we locate the receiving end through the triangulation algorithm. Our experimental results show that our positioning accuracy can reach 1.9 cm in consideration of the real-time performance of integrated positioning, which is sufficient for most indoor positioning applications.

References

[1] R. Zhang *et al.*, “Image sensor based visible light positioning system with improved positioning algorithm,” *IEEE Access*, vol. 5, pp. 6087–6094, 2017.
[2] M. Liu *et al.*, “Towards indoor localization using visible light communication for consumer electronic devices,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2014, pp. 143–148.
[3] H. T. Do and M. Yoo, “Visible light communication based vehicle positioning using a rolling shutter CMOS sensor,” in *Proc. 8th Int. Conf. Ubiquitous Future Netw.*, 2016, pp. 48–50.
[4] Y. J. Kim *et al.*, “High-resolution indoor positioning using light emitting diode visible light and camera image sensor,” *IET Optoelectron.*, vol. 10, no. 5, pp. 184–192, 2016.
[5] M. S. Hossen, Y. Park, and K. D. Kim, “Performance improvement of indoor positioning using light-emitting diodes and an image sensor for light-emitting diode communication,” *Opt. Eng.*, vol. 54, no. 3, 2015, Art. no. 035108.
[6] H. Huang *et al.*, “Indoor positioning method based on metameric white light sources and subpixels on a color image sensor,” *IEEE Photon. J.*, vol. 8, no. 6, Dec. 2016, Art. no. 6806110.

[7] R. Zhang *et al.*, “A single LED positioning system based on circle projection,” *IEEE Photon. J.*, vol. 9, no. 4, Aug. 2017, Art. no. 7905209.
[8] H. Phat and Y. Myungsik, “VLC-based positioning system for an indoor environment using an image sensor and an accelerometer sensor,” *Sensors*, vol. 16, no. 6, p. 783, 2016.
[9] R. Zhang *et al.*, “A novel sensor fusion based indoor visible light positioning system,” in *Proc. Int. Workshop GLOBE-COM*,” 2017, pp. 1–6.
[10] E. S. Maini, “Enhanced direct least square fitting of ellipses,” *Int. J. Pattern Recognit. Artif. Intell.*, vol. 20, no. 6, pp. 939–953, 2006.
[11] A. Baraldi and F. Parmiggiani, “Novel neural network model combining radial basis function, competitive Hebbian learning rule, and fuzzy simplified adaptive resonance theory,” *Proc. SPIE - Int. Soc. Opt. Eng.*, vol. 3165, pp. 98–112, 1997.
[12] G. Lera and M. Pinzolas, “Neighborhood based Levenberg–Marquardt algorithm for neural network training,” *IEEE Trans. Neural Netw.*, vol. 13, no. 5, pp. 1200–1203, Sep. 2002.
[13] C. Xie *et al.*, “The LED-ID detection and recognition method based on visible light positioning using proximity method [J],” *IEEE Photon. J.*, vol. 10, no. 2, Apr. 2018, Art. no. 7902116.