
# High precision indoor positioning method based on visible light communication using improved Camshift tracking algorithm

Hongyun Xie, Zequn Tan, Wanchao Mao, Zhipeng Liu, Manxi Liu

PII: S0030-4018(20)30194-2
DOI: https://doi.org/10.1016/j.optcom.2020.125599
Reference: OPTICS 125599

To appear in: *Optics Communications*

Received date : 7 January 2020
Revised date : 18 February 2020
Accepted date : 24 February 2020

Please cite this article as: H. Xie, Z. Tan, W. Mao et al., High precision indoor positioning method based on visible light communication using improved Camshift tracking algorithm, *Optics Communications* (2020), doi: https://doi.org/10.1016/j.optcom.2020.125599.

# High precision Indoor Positioning Method based on Visible Light Communication Using Improved Camshift Tracking Algorithm

**Hongyun Xie [a, †], Zequn Tan [b, †], Wanchao Mao[c,*], Zhipeng Liu [d,*], Manxi Liu [b]**

[a]Laboratory of Nuclear Power Safety Monitoring Technology and Equipment, China Nuclear Power Engineering Co., Ltd, Shenzhen, Guangdong, China, 518172

[b]South China University of Technology, School of Automation Science and Engineering, Guangzhou, Guangdong, China, 510640

[c]State Key Laboratory of Nuclear Power Safety Monitoring Technology and Equipment, Shenzhen, Guangdong, China, 518172

[d]South China University of Technology, School of Electronic and Information Engineering, Guangzhou, Guangdong, China, 510640

†Hongyun Xie, Zequn Tan are co-first authors of this article.

* Wanchao Mao, E-mail: maowanchao@cgnpc.com.cn; Zhipeng Liu, E-mail: scutlzpee@163.com

**Abstract**. Recently, visible light communication (VLC) has been widely used in indoor positioning, which has such advantages as cost-saving, environmentally friendly, anti-radio frequency interference (anti-RF-interference) and so on. Considering that existing VLC-based indoor positioning systems suffer from such obstacles as blur effects and shield effects, a precise robust positioning method using improved Continuously Adaptive Meanshift (Camshift) tracking algorithm is proposed to locate the positioning terminal. The improved Camshift algorithm is used to track the region of interest (ROI) of the LED to improve the robustness of the visible light positioning (VLP) system. Classical Camshift algorithm build an one-dimensional histogram with Hue component from HSV (hue, saturation, value) color space, which may cause tracking failure when illumination variation or similar color interferes occur. Therefore, the proposed algorithm utilizes hue and saturation components from HSV space to build 2-Demensional (2-D) color feature histogram. The H and S components of the HSV are combined with the Camshift algorithm to improve tracking accuracy and robustness. What's more, even most parts of light-emitting diode (LED) are shielded or broken, high-precision tracking can still be achieved by recognizing the color features and local detection regardless of shape change, thereby ensuring the positioning robustness. Furtherly, shield effects and background interferences are introduced to simulate actual positioning scenes. Experimental results show that the proposed algorithm can provide an average accuracy of 0.95 cm and ensure that 90% of total tracking error is less than 1.79 cm, indicating that the LED tracking is so accurate that the positioning accuracy of the positioning algorithm is not affected by the tracking algorithm. Meanwhile, the average computing time of the tracking algorithm is 0.036 s for per frame, which demonstrates that the computational cost required for Camshift is so small that the positioning algorithm is not affected by the tracking algorithm and can still have good real-time performance. Therefore, the proposed algorithm has broad application prospects in fields of dynamic positioning and tracking services.

**Keywords**: visible light communication (VLC); indoor positioning; improved Camshift algorithm; positioning accuracy; real-time ability; robustness; dynamic positioning and tracking.

## 1 Introduction

Nowadays, indoor positioning has become a research hotspot to meet increasing life needs and business demands, because it is vital for the tracking of mobile intelligent devices and navigation service. Traditional indoor positioning systems include wireless local area network (WLAN), Zig bee, Ultra Wideband (UWB), Bluetooth, Radio-frequency Identification (RFID), Infrared Ray and Ultrasonic Wave. However, UWB and RFID require additional infrastructure and new hardware components, and other RF-based systems fail to satisfy the requirements for positioning accuracy and real-time ability at the same time. Different from conventional indoor positioning technologies mentioned above, the visible light positioning (VLP) technology is a kind of indoor positioning technology based on visible light communication (VLC) technology[1,2]. By comparison, VLP has such advantages as cost-saving, environmentally friendly, anti-RF-interference and so on, which shows great commercial prospects. At present, VLP systems can be divided into two categories: photodiode-based (PD-based) and image sensor-based (IS-based)[3]. The PD-based VLP systems have been deeply studied by researchers before. And through our previous works exploring PD-based VLP systems[4-7], we found that due to PD's sensibility to the direction of the light beam, the mobility of location terminal suffers considerable restraints. Another deficiency of PD-based VLP systems is poor robustness, which is indicated by the light intensity fluctuation measured repeatedly at the same location. Besides, PD-based VLP systems are easily affected by ambient light and reflected light[8]. Meanwhile, precise measurement for angles and signal strength received are required, otherwise noticeable positioning error will appear. By contrast, the IS-based VLP systems are free from the problems mentioned above. Generally, smart phones are equipped with high-resolution complementary metal-oxide-semiconductor (CMOS) sensor cameras nowadays, so it can be easily combined with IS-based VLP methods without extra facility cost, which is of great commercial prospect. What's more, applications are restricted to low-speed motion or static positioning in most PD-based VLP systems, while dynamic positioning is seldom realized, which is significant for real-time indoor navigation[9]. In this sense, image sensor outperforms PD as the receiving terminal for VLP systems[10].

However, there exist two main problems in IS-based VLP systems: (1) Most systems fail to achieve satisfactory performance in terms of positioning accuracy, real-time ability and robustness, which are three essential factors in indoor positioning system. (2) Few VLP systems consider the case where LED is partially or completely shielded because of obstacles in way of propagation between the LED and positioning terminal. In[11], a circle-projection-based single LED positioning system is proposed, but its positioning accuracy is only 25.12 cm. It still remains to be improved in regard to positioning accuracy. In[12], a VLP method based on low-resolution photodiode array and front camera of mobile phone is proposed, whose 3-D positioning error is at the level of decimeter. In[13], electronic compass and the gyroscope sensor measure together with the scene image are utilized to realize the positioning of terminal within the accuracy of about 2 centimeters. However, the real-time positioning ability is ignored. The algorithm proposed in[14] achieves a high positioning accuracy at mm-level, but it only focuses on the static positioning and overlooks the real-time performance. In[15], Minimax filter is used to track for the terminal at an average velocity of 1 m/s (3.6 km/h) simulated by MATLAB simulator, which is unable to fit higher motion speed and still can't meet the requirement of real-time ability. In[16], the proposed VLP method allows the positioning terminal to reach a maximum motion speed of up to 18 km/h with a positioning accuracy of 7.5 cm. It strikes a good balance between accuracy and real-time ability, but it can't get rid of background interference caused by other luminous objects, for its positioning method is based on pixel intensity detection. When LED is disturbed by other luminous objects in the image, it becomes difficult to obtain the position of the LED through pixel intensity detection, which may lead to location failure. Besides the background interference, another fatal constraint to real-time ability and robustness of VLP is shielding effect. Because most of the positioning algorithms in VLP are based on 2 or more LEDs, when the optical links between the LEDs and the positioning terminal are blocked, the precise location would fail. But few existing works have considered such case when LED is partially or completely shielded.

To address the shortcomings of existing methods and improve the performance of our previous works, in this paper, we propose a precise high-speed robust positioning method using improved Camshift tracking algorithm based on 2-Demensional (2-D) color feature histogram, which has the following advantages: Firstly, the proposed Camshift algorithm utilizes color features to detect and track the positions of LED to further ensure tracking accuracy. Secondly, even most parts of the LED are shielded in the image, high-precision tracking can still be achieved by recognizing the color features regardless of shape change, which ensures the robustness of the VLP system without affecting the positioning accuracy. Thirdly, the proposed algorithm obtains the target position to the next frame by searching near the current frame target position. However, based on pixel intensity detection, the entire image is searched every time to get the position of the target, so our algorithm has less computational cost, which demonstrates that our proposed VLP system achieves high robustness without affecting the real-time ability of the positioning algorithm. Therefore, the proposed Camshift algorithm can accurately track the position of the LED in the image sequences in real time.

The remainder of this paper is organized as follows: Section 2 provides a detail description of the proposed positioning and tracking algorithm. Section 3 sets up the experiment and presents the analysis to verify the proposed algorithm. Section 4 gives the conclusion of the article.

## 2. Theory

*2.1 IS-based Visible Light Positioning Technology*

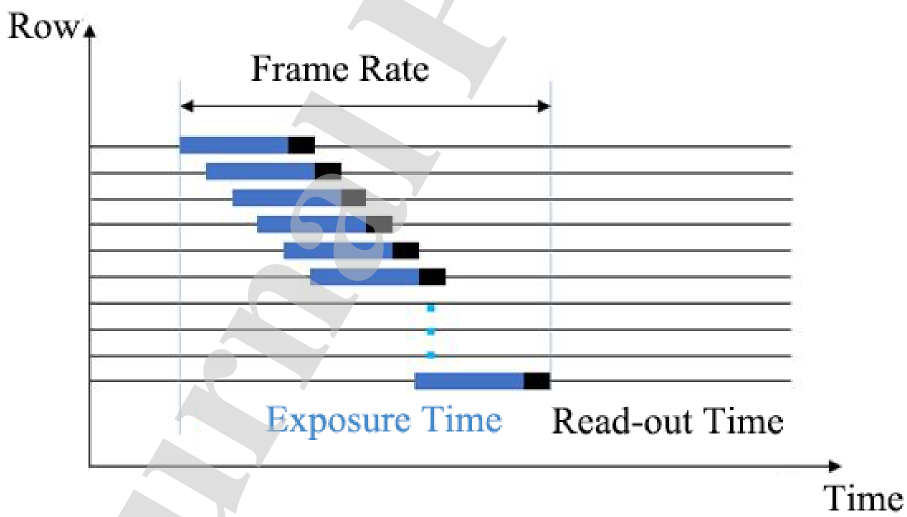



**Fig. 1** The rolling shutter mechanism of the CMOS.

The IS-based VLP realizes LED-ID recognition by utilizing the rolling shutter mechanism of the Complementary Metal Oxide Semiconductor (CMOS) image sensor. As shown in Fig. 1, the exposure and data readout are performed row by row, the data of one row read out immediately when the exposure of this row is finished. This is known as the rolling shutter mechanism. By the rolling shutter mechanism of CMOS sensor, turning on and off the LED light during a period of exposure would result in bright and dark stripes on the image captured by CMOS sensor. Unlike the traditional optical camera communication (OCC) method, we use the OOK-PWM modulation method to introduce different features (including frequency, duty cycle, phase difference, the area of the LED) to the light and dark stripes captured in the image for different

LED-ID, rather than directly demodulation. In our previous research [17,18], we have described the detailed processing of LED-ID modulation and features recognition.

In our VLP system, the implementation of dynamic positioning and tracking is to analyze and search the captured image sequence to find the position with the highest similarity to the target. After detecting the region of interest (ROI) of the LED in the initial image frame, the proposed Improved Camshift algorithm will estimate the ROI region of the LED in real time and dynamically in the subsequent frame. Estimating the ROI region of the LED means obtaining the position of the LED in the image sequence. In addition, after the LED is detected in the first frame, the LED-ID is identified by machine learning[18]. Finally, the positioning algorithm utilizes the obtained LED position to acquire the position of the camera. In the following sections, the single-lamp positioning technology is briefly introduced, and the Camshift algorithm is explained in detail in theory.

### *2.2 Single-lamp Positioning Technology*

Define $L^t$ as the set of pixels belonging to LED at time t in the image, and the pixel coordinates $\left(x_i^t, y_i^t\right) \in L^t$, where i represents ith pixel in the image[19]. In addition, the area of background is much larger than that of LED, so the position of LED in the image can be assumed as the center of $L^t$, which is defined as $\left(x_{i0}^t, y_{i0}^t\right)$.

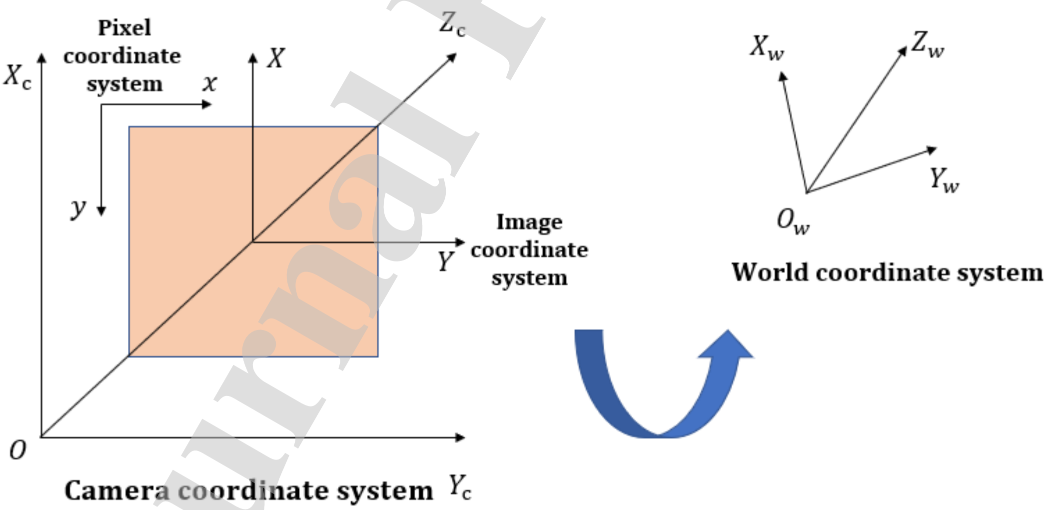



**Fig. 2** World coordinate system, Camera coordinate system, Image coordinate system and Pixel coordinate system

As shown in Fig.2, by utilizing the relationships between the various coordinate systems, the mapping from the pixel coordinate to the world coordinate can be established as follows:

Firstly, the image coordinate $\left(x^t, y^t\right)$ of the LED in the image is calculated as:

$$\begin{bmatrix} x_{i0}^t \\ y_{i0}^t \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & x_0 \\ 0 & \frac{1}{dy} & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^t \\ y^t \\ 1 \end{bmatrix} \qquad (1)$$

where $(x_0, y_0)$ represents the center of the image sensor imaging plane, i.e. the origin of image coordinate system. The unit of the pixel coordinate is pixel, which is described by its row and line. The unit of the image coordinate is mm, which belongs to physical unit. $\mathrm{dx}$ and $\mathrm{dy}$ represent the unit conversion of two coordinate system, i.e. 1 pixel= $\mathrm{dx}$ mm.

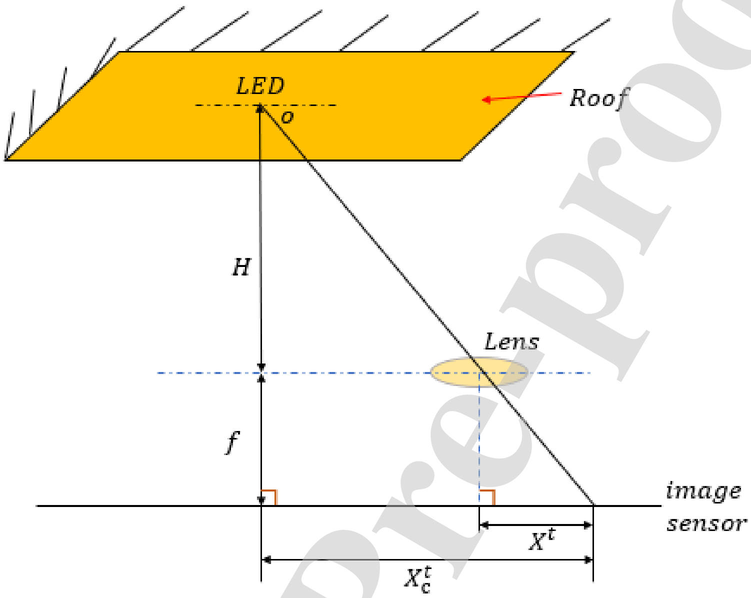



**Fig. 3** The geometrical relationship between the image coordinate system and the camera coordinate system

According to the geometric relationship between the image coordinate system and the camera coordinate system shown in Fig.3, the camera coordinate $(X_C^t, Y_C^t)$ of the LED can be computed as:

$$\frac{X_C^t}{X^t} = \frac{Y_C^t}{Y^t} = \frac{f + H}{f} \tag{2}$$

where f is the focus of lens, and H is the vertical distance between the LED and the lens plane.

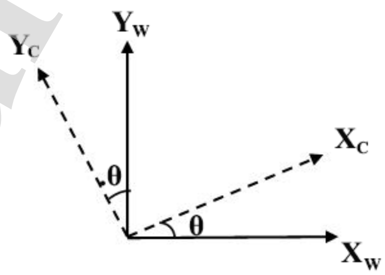



**Fig. 4.** The transformation model of camera coordinate system and the world coordinate system.

As can be seen from Fig.4, the camera coordinate can be rotated and translated to realize positioning under any azimuth. The camera coordinate $(X_C^t, Y_C^t)$ can be transformed to the world coordinate $(X_W^t, Y_W^t)$ of the image sensor through a matrix as follows:

$$\begin{bmatrix} X_C^t \\ Y_C^t \\ 1 \end{bmatrix} = \begin{bmatrix} cos\theta & sin\theta & 0 \\ -sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w^t \\ Y_w^t \\ 1 \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad (3)$$

where θ denotes the rotate angle between the world coordinate and the camera coordinate system and $[T_x \quad T_y \quad T_z]^T$ denotes translation.

*2.3 Improved Camshift Tracking Algorithm Based on 2-D Color Feature Histogram*

After the LED is detected using a LED detection algorithm in the first frame, the Improved Camshift algorithm establishes a probability distribution map of the target signal LED, and uses the Meanshift algorithm to iteratively search around the center of the target LED to obtain the target position of the next frame. In some traditional VLP methods based on image recognition only, every image must be processed to acquire the position of the LED, which results in large computational cost and poor real-time performance. However, if the features of the Camshift algorithm are utilized, only the local part of each image needs to be processed to obtain the position of the LED, which will greatly reduce the computational cost to ensure excellent real-time performance.

What's more, in the VLP system, the shielding effect and background interference are two obstacles that have the greatest impact on the robustness under actual conditions, which may lead to the positioning failure. However, many studies focus only on real-time but ignore robustness. In order to solve this problem, the Camshift algorithm achieves tracking LED in the shielding effect and background interference by establishing a probability distribution map of the target model and the candidate model.

*2.3.1 The Target Mode*

When the LED is detected in the first frame, the target model is established by weighting the HSV histogram using the kernel density estimation function Epannechnikov [20].

Assume that $\{x_1, \cdots, x_n\}$ is the pixels at different positions in the target LED region, then the probability density of the target area:

$$\hat{q}_u = C \sum_{i=1}^{n} K_E(\left||x_i|\right|^2)\delta(c(x_i) - u), u = 1, \cdots, m \quad (4)$$

where u is the index of the histogram; $c(x_i)$ indicates the pixel value of $x_i$ in the target area is mapped to the index value in the corresponding histogram; $\delta(.)$ is Kronecker delta function; C is the normalization constant, which is equal to $[\sum_{i=1}^{n} K_E(\left||x_i|\right|^2)]^{-1}$ and makes $\sum_{i=1}^{n} \hat{q}_u = 1$; $K_E(.)$ is the kernel density estimation function Epannechnikov:

$$K_E(x) = \begin{cases} c\left(1 - \left||x|\right|^2\right), \left||x|\right| \leq 1 \\ 0, otherwise \end{cases} \quad (5)$$

where x denotes the pixel position of the LED region.

*2.3.2. The Candidate Region Model*

Suppose $\{y_1, \cdots, y_{n_k}\}$ is the $n_k$ pixel points of the candidate target region, h is the number of the pixels of candidate target and the center of the current frame is y, the probability function of the candidate target is:

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n} K_E\left(\left\|\frac{y-y_i}{h}\right\|^2\right)\delta(c(y_i)-u), u = 1, \cdots, m \tag{6}$$

$C_h$ is the normalization constant, making $\sum_{i=1}^{n} \hat{p}_u = 1$:

$$C_h = \left[\sum_{i=1}^{n_k} K_E\left(\left\|\frac{y-y_i}{h}\right\|^2\right)\right]^{-1} \tag{7}$$

*2.3.3. Similarity Measure*

After obtaining the probability function of the target model and the candidate target respectively, the Bhattacharyya distance[21] can be used to calculate their similarity between $\hat{q}_u(x)$ and $\hat{p}_u(y)$:

$$d(y) = \sqrt{1-\hat{\rho}(y)} \tag{8}$$

$$\hat{\rho}(y) = \rho[\hat{q}, \hat{p}(y)] = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)\hat{q}_u} \tag{9}$$

Where $\hat{\rho}(y)$ is the Bhatacharyya (BH) coefficient between $\hat{q}$ and $\hat{p}(y)$, which belongs to [0,1]. If its value is closer to 1, the similarity between the candidate region model and the target model is, the closer it is to the true position of the target LED.

*2.3.4. Meanshift Algorithm*

In order to find the candidate model that matches the target model the most, the value of BH must be maximized[22].

Firstly, the target center $y_0$ of the previous frame is used as the initial position of the current frame search area and the color feature probability density function $\hat{q}_u$ is calculated.

Secondly, perform iterative search near $y_0$ to obtain the new position $y_1$:

$$y_1 = \frac{\sum_{i=1}^{n_k} x_i \omega_i}{\sum_{i=1}^{n_k} \omega_i} \tag{10}$$

where $\omega_i = \sum_{u=1}^{m} \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}}\delta(c(x_i)-u)$ represents kernel weights.

Thirdly, calculate the $\hat{p}_u(y_1)$ and BH coefficient $\rho[\hat{q}, \hat{p}(y_1)]$ at the new position $y_1$. Then, judge if $\rho[\hat{q}, \hat{p}(y_1)] < \rho[\hat{q}, \hat{p}(y_0)]$ is true. If it is true, assign $\frac{1}{2}(y_0 + y_1)$ to $y_1$.

Finally, judge whether $||y_1 - y_0|| < \varepsilon$ is true. If it is true, terminate the loop judgment. At this time, $y_1$ is the target position. If not, the value of $y_1$ is assigned to $y_0$, and then the loop calculation is performed according to the formula (9).

### *2.3.5. HSV Histogram and Back Projection*

Traditional Camshift algorithm uses Hue channel in HSV space to construct histogram, so when Saturation or Value vary greatly, its tracking sensitivity decreases significantly. Therefore, the basic concept of our proposed algorithm is to introduce hue and saturation components from HSV space to obtain 2-D color joint histogram, according to which the original image is transformed into color probability distribution image.

In the 2-D histogram, hue component ranges from $0^{\circ}$ to $360^{\circ}$ and saturation component ranges from 0 to 1. For convenience, the range of Hue component is quantified as 30 bins with a quantified step of $12^{\circ}$, and the range of Saturation component is quantified as 16 bins with a quantified step of $\frac{1}{12}$.

According to the obtained color histogram, the original image is transformed into a color probability distribution image, which is called "Back Projection". The mentioned Histogram represents the probability or pixel number of different Hue or Saturation component values. It is equivalent to a color probability lookup table. Define C(h) as the number of pixels whose H component is h in the histogram and max [C(h)] denote the maximum value of C(h). Let C(s) denote the number of pixels whose S component is s in the histogram and max[C(s)] denote the maximum value of C(s). Then the probability $p(h_0, s_0)$ that the pixel whose H component is $h_0$ and S component is $s_0$ is the image color is:

$$p(h_0, s_0) = \frac{C(h_0) \times C(s_0)}{\max[C(h)] \times \max[C(s)]} \tag{11}$$

Then the corresponding relationship is established between H， S component and image color probability, which serves as an image color probability lookup table. By replacing the value of each pixel in the image with the appearing probability of its color, the color probability distribution image is obtained, which is a gray image.

### *2.3.6. Adaptive Adjustment of The Search Window*

After obtaining the initial position and size information of the target object in the first frame according to the initial information of image and search window, the information of current frame is transmitted to the next frame as the initial value. In the video sequence acquired by camera in real time, the process above is repeated to continuously track for target objects. The proposed Camshift algorithm obtains the information of initial search window for the target object, including the size of the search window and position of its centroid. Then, according to the movement of the target object, the search window is updated continuously, so that it can realize the detection and recognition of the continuous moving object[23].

After select a search window W in color probability distribution image, Camshift algorithm adjusts the search window in real time, and keeps updating the centroid of the search window, which needs to calculate zero-order moment $M_{00}$, first-order moment $M_{10}, M_{01}$ and second-order moment $M_{20}, M_{02}, M_{11}$ as follows:

$$M_{ij} = \sum_x \sum_y x^i y^j I(x,y) \tag{12}$$

where $i = 0,1,2; j = 0,1,2; I(x,y)$ is the probability value at the coordinate $(x,y)$ in the probability distribution diagram. The centroid position $(x_c, y_c)$ of search window in current frame is represented as:

$$x_c = \frac{M_{10}}{M_{00}} \tag{13}$$

$$y_c = \frac{M_{01}}{M_{00}} \tag{14}$$

The result $(x_c, y_c)$ is viewed as the center of search window for next frame. With the change of the target, the centroid information of the search window is updated. By adjusting the size of the search window, the ellipse information of the search window including the length of long axis $l$, the length of short axis $w$ and direction angle $\theta$ are represented as:

$$l = \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}} \tag{15}$$

$$w = \sqrt{\frac{(a+c) - \sqrt{b^2 + (a-c)^2}}{2}} \tag{16}$$

$$\theta = \tan^{-1}\left(\frac{b}{a-c}\right) \tag{17}$$

where a,b,c can be calculated as:

$$a = \frac{M_{20}}{M_{11}} - x_c^2 \tag{18}$$

$$b = 2\left(\frac{M_{11}}{M_{00}} - x_c y_c\right) \tag{19}$$

$$c = \frac{M_{02}}{M_{11}} - y_c^2 \tag{20}$$

As can be seen from the formulas above, zero-order moments represent the sum of the pixels in the search window, which represents the size of the search window. When the calculated zero-order moment is greater than the set threshold, it shows that the current target object is full of the whole search window, and it is necessary to enlarge the search window to make the target object included by the camera's field of vision. Conversely, when the threshold is larger, the size of the search window should be reduced accordingly.

*2.3.7. The Implementation of the Improved Camshift Algorithm*

The Improved Camshift tracking algorithm tracks the color information of the target object in video images. It uses the color histogram of target to transform the image into a color probability distribution image. Based on this idea, the size and position of search window is initialized. And the position and size of search window in the subsequent frame are adaptively adjusted according to the results of the previous frame to locate the central position of the target in current image. The adaptive search process of the Camshift algorithm is shown in Fig.5 and the specific implementation steps of the algorithm are as follows:

- Step 1 initialization:

Initialize the size of the search window $S_0$ and the initial position $(x_0, y_0)$ in the color probability distribution map.

- Step 2 establish HSV histogram:

Extract the H component and S component of the region where the target LED and establish the 2-D histogram.

- Step 3 Back projection:

If the current frame image is not empty, calculate its color probability distribution map (i.e. back projection).

- Step 4 Tracking target LED:

According to the back projection and the search window, the Meanshift algorithm is called to obtain a convergence window.

The Camshift algorithm adjusts the convergence window by the target orientation information reflected by the convergence window obtained by the Meanshift iteration, thereby determining the initial size of the search window when tracking the target in the next frame image.

- Step 5 Loop:

In the next frame image, the adjusted convergence window obtained in step 4 is used as the search window, and jumps to step 3 to continue execution.

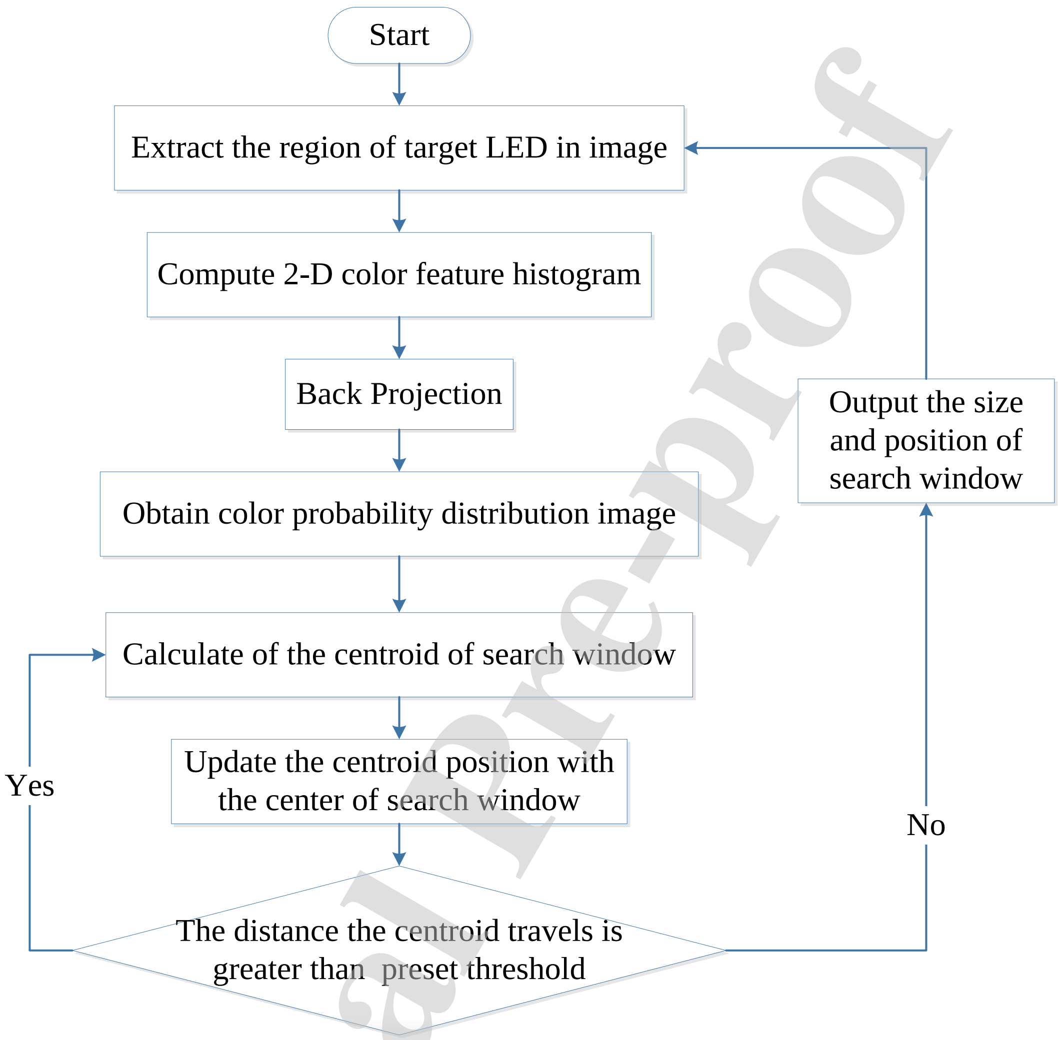



**Fig. 5** The flow diagram of the proposed improved Camshift algorithm.

## 3. Experiment Setup and Result Analysis

The experimental platform shown in Fig.6 consists of a constant voltage source, six LEDs, an industry camera, a smart car and a computer (Dell Inspiron 5557, Windows 10, 4G RAM, Intel (R) Core (TM) i5-6200 CPU @ 2.4GHz). Constant voltage source supplies power to LEDs. The LEDs as transmitter transmit modulated optical signals, which are captured by the industry camera mounted on a smart car as receiver. The smart car plays the role of indoor mobile intelligent devices, and can move along random track. Meanwhile, the computer processes receiving image signals with proposed algorithm to detect the LEDs and position the camera in real time. In our experiment, open source computer vision library (OpenCV 3.2.0) is utilized to process receiving images with C++ language as the software system. As can be seen from Fig.6, six LEDs are uniformly distributed on the ceiling of the system configuration in an experimental platform. Besides, specifications for the industry camera, experimental platform, LEDs, the smart car and the circuit board are shown detailly in Table 1.

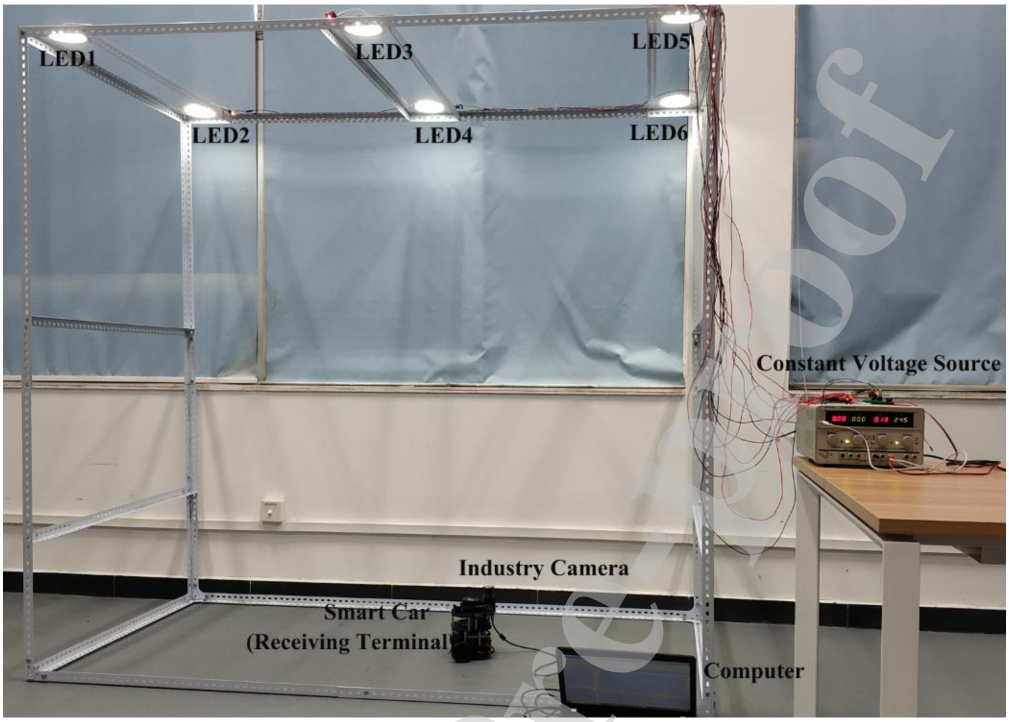



**Fig. 6** Experimental platform and hardware devices.

**Table 1.** Parameters of the experimental platform

| | |
|---|---|
| **Camera Specifications** | |
| Model | MV-U300 |
| Spectral Response Range/nm | 400~1030 |
| Resolution | 800×600 |
| Frame Rate/FPS | 46 |
| Dynamic Range/dB | >61 |
| Signal-to-noise Ratio/dB | 43 |
| Pixel(H×V) | 2048×1536 |
| Pixel Size/μm² | 3.2×3.2 |
| Time of Exposure/ms | 0.0556-683.8 |
| Sensitivity | 1.0V/lux-sec 550nm |
| Optical Filter | 650nm Low Pass Optical Filter |
| Type of Shutter | Electronic Rolling Shutter |
| Acquisition Mode | Successive and Soft Trigger |
| Working Temperatures/°C | 0-50 |
| Support Multiple Visual Software | OpenCV, LabView |
| Support Multiple Systems | Vista, Win7, Win8, Win10 |
| **Experimental Platform Specifications** | |
| Size (L×W×H)/ $cm^3$ | 210×110×200 |
| **LED Specifications** | |
| Coordinates (x, y, z)/cm | LED1(20, 20, 300)<br>LED2(200, 20, 300)<br>LED3(20, 150, 300) |

| | |
|---|---|
| | LED4(200, 150, 300)<br>LED5(20, 300, 300)<br>LED6(200, 300, 300) |
| Diameter of each LED/cm | 15 |
| Power of each LED/W | 6 |
| The half-power angles of LED/deg($\varphi_{1/2}$) | 60 |
| **Smart Car Specifications** | |
| Single-chip | STC89C52RC |
| Weight/g | 1200 |
| Remote Control | Bluetooth, Wifi |
| **Circuit Board Specifications** | |
| Drive chip | DD311 |
| Drive current/A | 0.1 |
| Drive voltage/V | 28 |

To better demonstrate the experimental results, we recorded a video sequence with the industry camera to verify the proposed algorithm. In the experiment, the smart car as receiving terminal keeps moving throughout the process, whose trajectory is controlled by a mobile phone through Bluetooth. With proper exposure time of the camera, only LEDs and other luminous objects are left in the images. Here, other luminous objects are regarded as background interferences. Besides, the single-lamp positioning technology is introduced into our VLP systems. After applying the relationships among multiple coordinate systems, the ROI region of camera shrinks to a small region so that we can realize positioning with the coordinate of only one LED (called target LED). In other words, there is no need to get the coordinates of all LEDs for positioning. As long as one of LEDs is obtained, the terminal equipment can be located, which decreases the computational cost. Therefore, to simplify the experiment, only one LED is taken into consideration to validate the algorithm performance.

### *3.1. Positioning Robustness*

Robustness refers to the ability of the VLP system to work properly under abnormal conditions, which is a vital indicator of positioning performance. In practical applications, the dynamic tracking and positioning systems based on visual VLC suffer from critical obstacles such as shielding effect and background interferences, which may lead to positioning failure. Shielding effect means the target LED is partially or completely blocked by obstacles on the propagation path. In addition, background interferences occur when other luminous objects apart from target LEDs are captured by the camera's field of view. These disturbances conduct a bad impact on tracking results and even prevent the ROI of the target LED in images from being extracted, causing the failure of LED-ID recognition. Therefore, these two situations are simulated to test the robustness performance of the proposed algorithm.

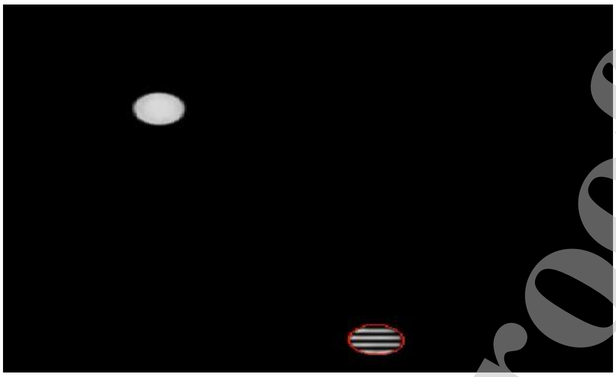
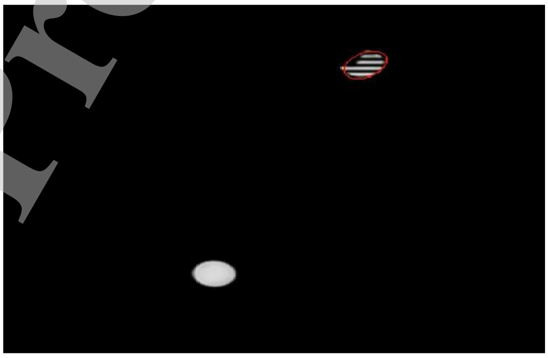


**Fig. 7** Complete signal source detection free from shield effects.

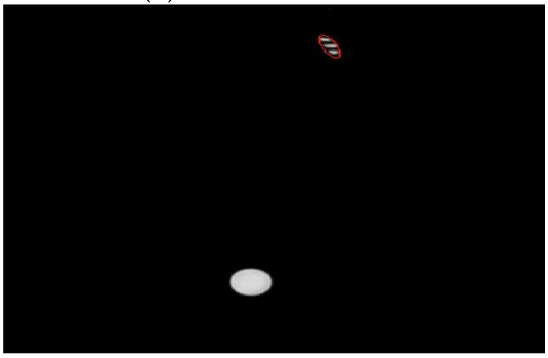
In most cases, the signal source appears completely in the camera's field of view nearly without being shielded, and the region of the target LED is extracted by the red search window accurately and fully. When there is a luminous object without modulation serves as background interference, it can be seen from Fig.7 that the target LED can still be accurately tracked by the proposed algorithm.

(**a**) frame number:192



(**b**) frame number: 223

(c) frame number:244



**(d)** frame number: 247

**Fig. 8** Tracking performances under shield effects of varying degrees.

In addition, when the target LED is shielded by different degrees, the extraction of the region of signal source and positioning are affected accordingly. In our experiment, the target LED is shielded to varying degrees from the 188th frame to the 273th frame. Fig.8 shows that the proposed algorithm performs well even under shield effects, and the target LED can still be tracked by the search window successfully. Besides, if the height changes, the target LED will become larger or smaller in the image sequence. Providing that the search window is fixed and the area of the LED in the image is larger than that of the search window at this time when the height changes, the ROI area of the LED obtained by the search window may be only a corner of the LED light. The position of the LED obtained by tracking at this time is obviously inaccurate. However, since the Camshift algorithm can automatically adjust the size of the red according to the change of the target shape, it can extract the target LED area more accurately, which is more suitable for real life applications.

The above experiments prove that the proposed algorithm can still work normally with interferences. Therefore, the proposed algorithm has satisfactory performance in positioning robustness.

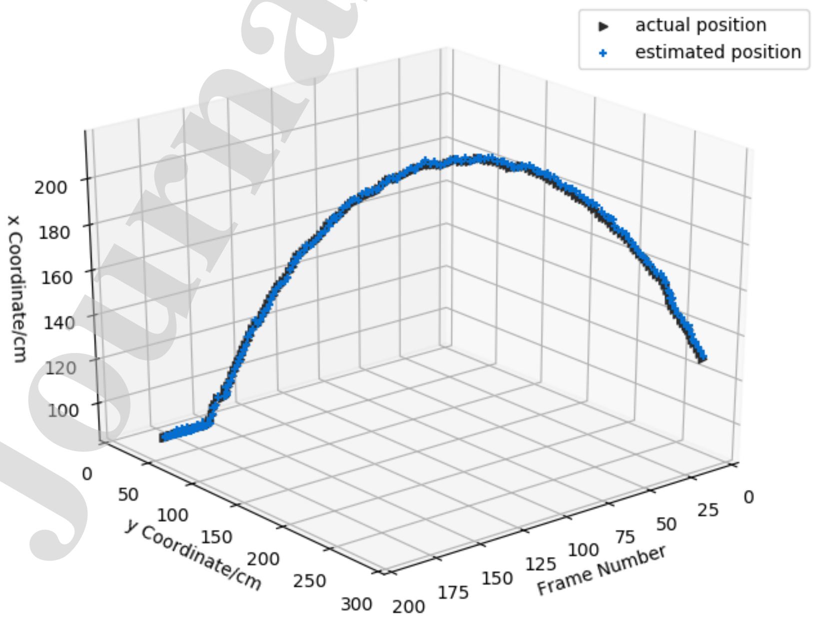
*3.2. Tracking Accuracy*

Tracking accuracy is the major evaluating index of positioning performance, and positioning error D is defined as:

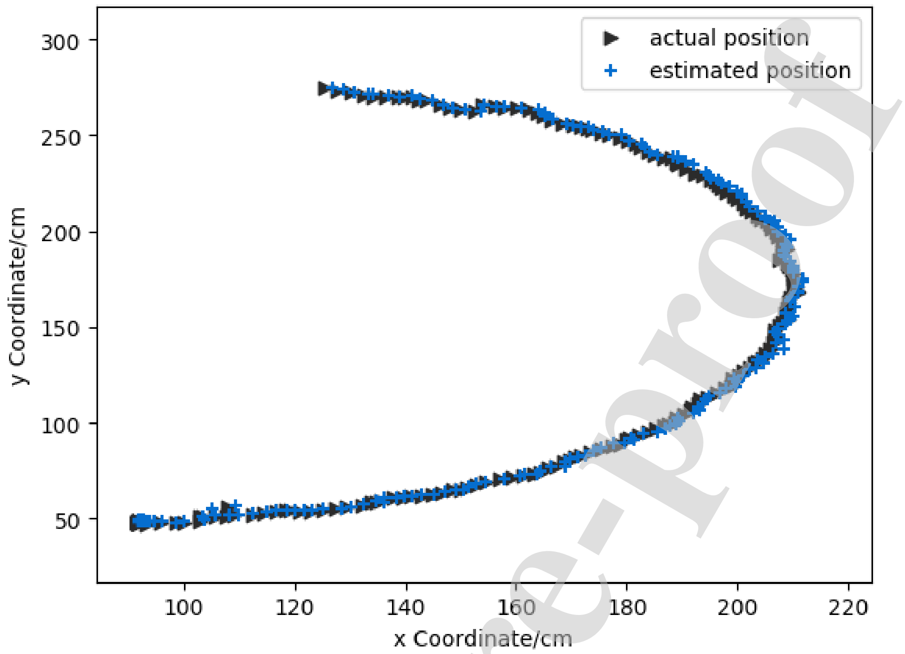
$$D = \sqrt{(x_a - x_e)^2 + (y_a - y_e)^2} \tag{21}$$

$$error\ of\ x_a(y_a) = |Value_{calculated} - Value_{actual}| \tag{22}$$

where $(x_a, y_a)$ is the estimated position calculated by the proposed algorithm, and $(x_e, y_e)$ represents the actual position. In the recorded video sequences, 194 consecutive frames (from the 100th frame to the 293th frames) are selected to evaluate the accuracy performance of the proposed algorithm. The tracking result is shown in Fig.9, where the blue plus sign is the coordinates predicted by the tracking algorithm and the black arrow is the real coordinate. As can be seen intuitively from the result, the blue plus and the black arrow almost overlap, and the tracking performance is excellent for each sample with small error. To better illustrate the result and validate the performance of the proposed algorithm, Fig.10 shows the tracking error of x and y coordinates respectively. The average error of x coordinate is 0.72 cm and average error of y coordinate is 0.48 cm.




**(a)**




**(b)**

**Fig. 9** Tracking results of the proposed algorithm.

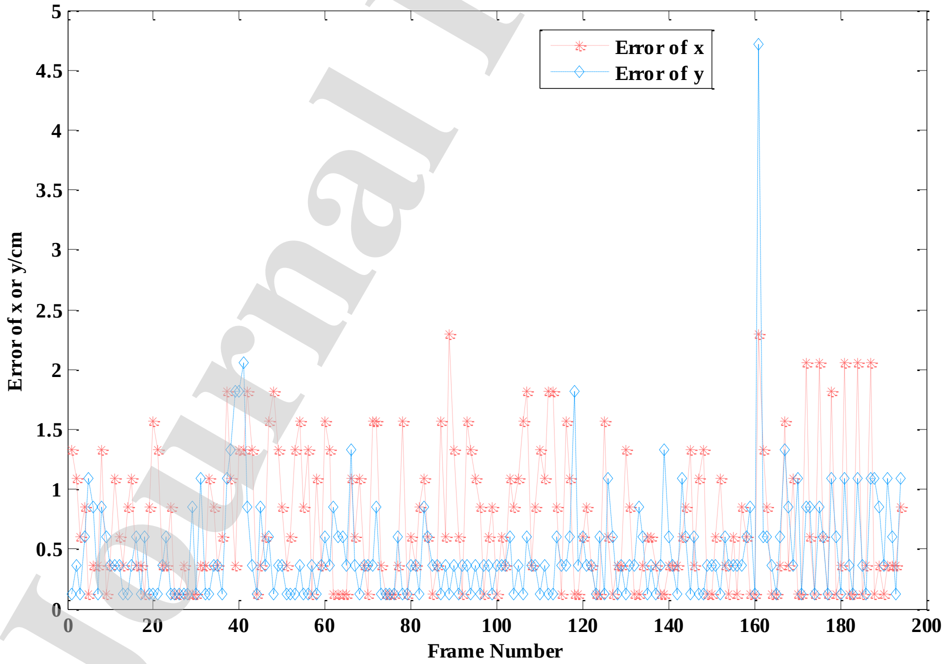



**Fig. 10** Tracking error of x and y coordinates.

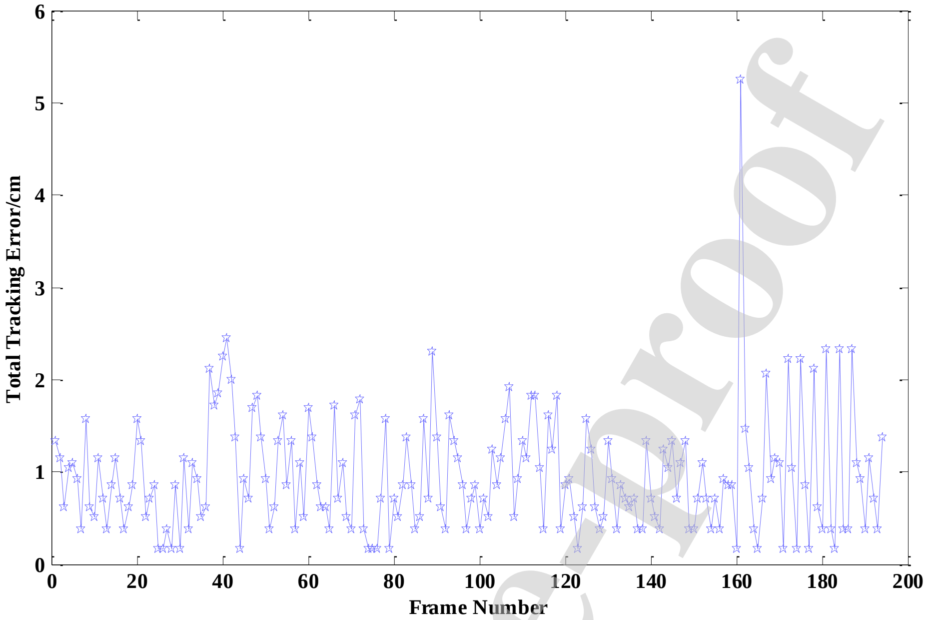



**Fig. 11** Total tracking error of the proposed algorithm.

Fig.11 shows the tracking errors of different frames in recorded video sequences. As can be seen, the maximum tracking error is 5.25 cm, and average tracking error is 0.95 cm, which denotes that the LED detection is so accurate that the positioning algorithm is not affected by the tracking algorithm. If the tracking algorithm has a large tracking error on the LED position, the positioning accuracy acquired by the positioning algorithm using the calculated LED position will be greatly reduced in the camera positioning. In addition, the average error from the 188th frame to the 273th frame where the target LED is shielded to varying degrees is 2.69 cm. Namely, the average tracking error under shield effect is 2.69 cm. Besides the shielding effects, there are still many causes of the tracking error, such as image noises, various background interferences in real scene, and so on. However, in most practical applications of dynamic positioning and tracking systems, such error is negligible compared with the size of positioning terminal, so a cm-level accuracy is acceptable. Therefore, the proposed algorithm can conduct a pleasing tracking and positioning performance.

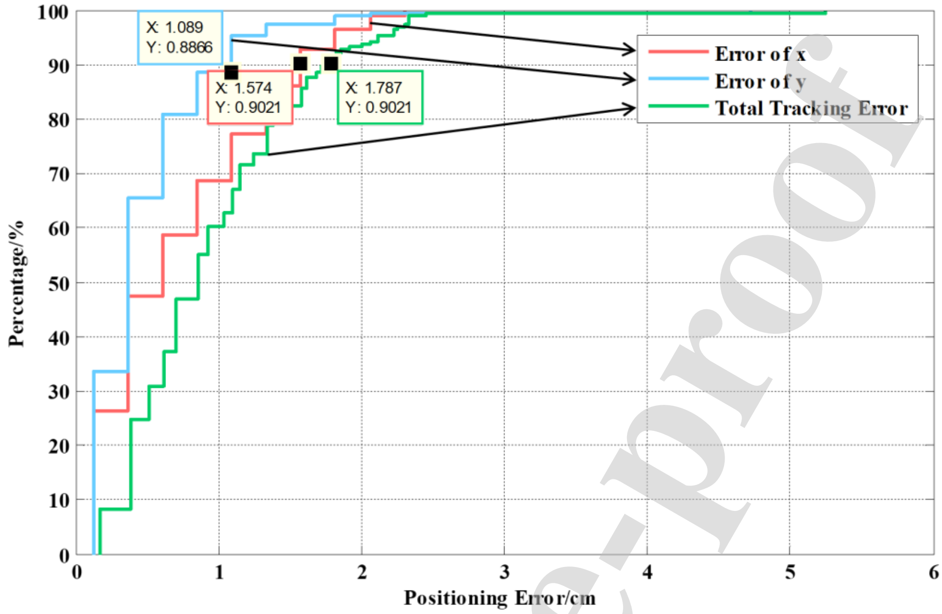



**Fig. 12** The CDF curves of tracking error.

Furthermore, the cumulative distribution function (CDF) curves of tracking errors in different dimensions are drawn in Fig.12. As indicated by the curves, more than 90% of total tracking error is less than 1.79 cm, more than 90% of error of x is less than 1.58 cm, and more than 88% of error of y is less than 1.09 cm. If 90% is assumed as an acceptable service coverage rate, the proposed algorithm is capable of delivering an accuracy of 1.79 cm. Therefore, it can be concluded that the proposed algorithm delivers a good tracking accuracy performance.

### *3.3. Real-time Ability*

In dynamic tracking and positioning systems, real-time ability is crucial, otherwise real-time tracking and positioning will fail. Since the terminal keeps moving throughout the process, when the estimated coordinate is calculated, the positioning terminal has already moved to the next position. If the algorithm cannot catch up in time, the position deviation will lead to noticeable positioning error. However, providing that the processing time of the program is short enough, such errors will not have much adverse effect on the implementation of tracking positioning in most applications.

In our experiment, 194 successive frames are chosen and their average processing time is 0.036 s for per frame, which indicates that the Camshift algorithm program is running for a short enough time that the positioning algorithm is basically not affected by the tracking algorithm and the VLP system can still have good real-time performance. Assume that the tracking algorithm takes 10s to obtain the position of the LED, the terminal position obtained by the positioning algorithm is the terminal position before 10s, demonstrating that the VLP system cannot achieve real-time tracking and positioning, which cannot be applied in real life. Compared to our previous works in the field of dynamic tracking and positioning, the proposed algorithm in this paper is more high-speed. In our previous work [19], the accuracy of optical flow detection and Bayesian forecast algorithm is 0.86cm, which is similar to the accuracy of the Camshift algorithm. However, its average running time is 0.162s in[19], which is about four times as long as our computing time. In[24], the processing time of particle filter is 0.026s, which is similar to the running time of the Camshift algorithm, but its tracking accuracy can only reach 2.95cm, which is about 3 times of our tracking accuracy. Therefore, it can be concluded that the proposed algorithm behaves well in real-time performance and tracking accuracy.

## 4. Conclusions

In this paper, a real-time dynamic positioning and tracking method with high accuracy and strong robustness is proposed, which is based on VLC-based indoor positioning system and transformations among different coordinate systems. The proposed method utilizes improved Camshift algorithm based on 2-D color feature histogram to track the ROI region of the LED, which solves two major problems in VLP systems: background interference and shielding effect. The proposed algorithm obtains the target position to the next frame by recognizing the color features and searching near the current frame target position, which effectively solves the problem of background interference and shielding effect while ensuring the robustness and the real-time performance of the VLP system without affect the positioning accuracy and the real-time ability of the positioning algorithm. Besides, both H component and S component of HSV histogram are considered to improve real-time ability and robustness, making it more feasible in a practical scenario with various interferences.

In our experiments, shielding effects and background interferences are introduced to simulate actual positioning scenes. Our experimental results show that the proposed algorithm can provide an average accuracy of 0.95 cm and an average computing time of 0.036 s for per frame, proving that the proposed algorithm can improve the robustness of the system without affecting the positioning algorithm. The combination of the improved Camshift algorithm and the positioning algorithm forms the VLP system that can possess a good performance in terms of positioning accuracy, real-time ability and robustness. Therefore, the proposed VLP method has broad application prospects in fields of dynamic positioning and tracking services.

*Acknowledgments:*

This work supported by National Undergraduate Innovative and Entrepreneurial Training Program (No.201510561003, 201610561065, 201610561068, 201710561006, 201710561054, 201710561057, 201710561058, 201710561199, 201710561202, 201810561217, 201810561195, 201810561218, 201810561219), Provincial Undergraduate Innovative and Entrepreneurial Training Program (NO. 201910561205), Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation (“Climbing Program” Special Funds) (pdjh2017b0040, pdjha0028, pdjh2019b0037), Guangdong science and technology project (2017B010114001).

*References*

1. Yasir Muhammad, Ho Siu-Wai, and Vellambi Badri N, “Indoor positioning system using visible light and accelerometer,” J. Lightwave Technol., vol. 32, no. 19, 2014, pp.3306- 3316.
2. Weizhi Zhang, Chowdhury M.I.Sakib, and Kavehrad Mohsen, “Asynchronous indoor positioning system based on visible light communications,” Opt. Eng., vol. 53, no. 4, 2014, Art. no.045105.
3. Do Trong-Hop and Yoo Myungsik, “An in-depth survey of visible light communication based positioning systems,” Sensors. J., vol. 16, no. 5, 2016, Art. no. E678.

4. Q. Peng, W. Guan, Y. Wu, Y. Cai, C. Xie, and P. Wang, “Three-dimensional high-precision indoor positioning strategy using Tabu search based on visible light communication,” Opt. Eng., vol. 57, no. 1, Jan. 2018, Art. no. 016101.
5. H. Chen, W. Guan, S. Li, and Y. Wu, “Indoor high precision three-dimensional positioning system based on visible light communication using modified genetic algorithm,” Opt. Commun., vol. 413, pp. 103–120, Apr. 15, 2018. Vol. 10, No. 3, June 2018.
6. W. Guan et al., “A novel three-dimensional indoor positioning algorithm design based on visible light communication,” Opt. Commun., vol. 392, pp. 282–293, Jun. 1, 2017.
7. Y. Cai, W. Guan, Y. Wu, C. Xie, Y. Chen, and L. Fang, “Indoor high precision three-dimensional positioning system based on visible light communication using particle swarm optimization,” IEEE Photon. J., vol. 9, no. 6, Nov. 11, 2017, Art. no. 7908120.
8. W. Guan, Y. Wu, C. Xie, H. Chen, Y. Cai, and Y. Chen, “High-precision approach to localization scheme of visible light communication based on artificial neural networks and modified genetic algorithms,” Opt. Eng., vol. 56, no. 10, Oct. 1, 2017, Art. no. 106103.
9. M. Yasir, S.-W. Ho, and B. N. Vellambi, “Indoor position tracking using multiple optical receivers,” J. Lightw. Technol., vol. 34, no. 4, pp. 1166–1176, Feb. 15, 2016.
10. W. Guan, Y. Wu, C. Xie, L. Fang, X. Liu, and Y. Chen, “Performance analysis and enhancement for visible light communication using CMOS sensors,” Opt. Commun., vol. 410, pp. 531–545, Mar. 1, 2018
11. R. Zhang, W.-D. Zhong, Q. Kian, and S. Zhang, “A single LED positioning system based on circle projection,” IEEE Photon. J., vol. 9, no. 4, Aug. 2017, Art. no. 7905209.
12. Y. Wang, Y. Gong, and Z. Shi, “Research on the collinear equation model of visual positioning based on visible light communication,” in Proc. MATEC Web Conf. Int. Conf. Eng. Technol. Appl., vol. 22, Jul. 9, 2015, pp. 02003-1–02003-6.
13. Ren Z.X., Zhang H.M., Wei L., et al. “A high precision indoor positioning system based on VLC and smart handheld [C],” Proceedings of the Computers and Information Processing Technologies I, 2014: 183-186
14. Md. 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, Mar. 1, 2015, Art. no. 035108.
15. M. Alsalami Farah, “Game theory Minimax filter design for indoor positioning and tracking system using visible light communications,” in Proc. 6th Int. Conf. Inf. Commun.23 Manage., Dec. 14, 2016, pp. 197–200.
16. J. Fang et al., “High-speed indoor navigation system based on visible light and mobile phone,” IEEE Photon. J., vol. 9, no. 2, Apr. 2017, Art. no. 8200711.
17. W. Guan, Y. Wu, C. Xie, L. Fang, X. Liu, and Y. Chen, “Performance analysis and enhancement for visible light communication using CMOS sensors,” Opt. Commun. J., vol. 410, pp. 531–545, Mar. 1, 2018.
18. C. Xie, W. Guan, Y. Wu, L. Fang, and Y. Cai, “The LED-ID detection and recognition method baed on visible light positioning using proximity method,” IEEE Photon. J., vol. 10, no. 2, Apr. 2018, Art. no. 7902116.
19. W. Guan, X. Chen, M. Huang, Z. Liu, Y. Wu, Y. Chen, “High-Speed Robust Dynamic Positioning and Tracking Method Based on Visual Visible Light Communication Using Optical Flow Detection and Bayesian Forecast,” IEEE Photonics Journal, vol. 10, no. 3, June 2018.
20. Zhao Q, Tao H, “Object Tracking Using Color Correlogram,” 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance. C., pp. 263-270, 2005.
21. D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. IEEE Trans. Pattern Anal. Mach. Intell. J., vol. 25, no.5, pp. 564-577, May 2003.

22. Comaniciu D, Ramesh V Meer P, "Real-time tracking of non-rigid objects using mean shift", IEEE Conference on Computer Vision and Pattern Recognition. C., vol. 2, pp. 142-149, 2000.
23. Bradski G.R. "Real Time Face and Object Tracking as a Component of a Perceptual User Interface," the Fourth IEEE Workshop on Applications of Computer Vision. C., pp. 214-219, 1998.
24. Y. Wu, W. Guan, X. Zhang, M. Huang, J. Cao, "Visible light positioning system based on CMOS image sensor using particle filter tracking and detecting algorithm", Optics Commun. J., vol. 444, pp. 9-20, August 2019.

Credit Author Statement



Hongyun Xie : Conceptualization; Zequn Tan: Writing - Review & Editing; Wanchao Mao: Methodology; Zhipeng Liu: Writing - Review & Editing; Manxi Liu: Formal analysis