

High Accuracy, 6-DoF Simultaneous Localization and Calibration Using Visible Light Positioning

Junye Chen , Dinghao Zeng, Chen Yang, and Weipeng Guan 

Abstract—Benefiting from the development of image sensors and the popularity of light-emitting diode (LED) lighting technology, visible light positioning (VLP) technology based on image sensors has ushered in vigorous development and broad prospects, which can provide low-cost and high-accuracy position service. However, the existing approaches require dense LEDs or sensors such as gyroscopes to assist positioning, which limits the area and lower the accuracy of positioning because of the errors from imperfect sensors. In this article, we propose a simultaneous localization and calibration VLP method based on double coplanar circular LED lights aiming to get rid of the dependence on additional sensors and dense LED transmitters. By the pinhole camera model and the perspective projection of circle, our proposed method extends the available position area and relaxes the required quantity of LED to two. The experiment result showed that our system has a mean 3D positioning accuracy of 7.91cm, a mean angle error of less than 1.6°, and an average latency of 182ms on mobile devices.

Index Terms—6 Degrees of freedom (6-DoF), circular projection features, high accuracy, visible light positioning (VLP).

I. INTRODUCTION

FOR years, positioning technologies have played an important role in many fields, such as transportation, production, navigation, and commercial service. When considering outdoor scenarios, wireless signals can travel in open spaces, so the problem is usually solved by technologies such as Global Positioning System (GPS). However, due to the obstruction of signal transmission, the performance of GPS decreases significantly when used indoors. Indoor positioning is widely used in shopping malls, museums, warehouses, parking lots, and other indoor spaces, making the demand for low-cost and high-accuracy positioning technology increase rapidly. However, most indoor positioning approaches, such as wireless fidelity (Wi-Fi), Zigbee, Bluetooth, Ultra-Wideband (UWB), Infrared Radiation (IR), ultrasonic and Radio-frequency Identification (RFID), are limited by either high cost when considering a large-scale deployment, or low accuracy typically in some scenarios.

Manuscript received 6 February 2022; revised 23 May 2022 and 3 July 2022; accepted 9 August 2022. Date of publication 16 August 2022; date of current version 4 November 2022. This work was supported by the National Undergraduate Innovative and Entrepreneurial Training Program under Grant 2021110561163. (Corresponding author: Weipeng Guan.)

The authors are with the School of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: 201930360107@mail.scut.edu.cn; 201830052058@mail.scut.edu.cn; 201930364105@mail.scut.edu.cn; augwpscut@mail.scut.edu.cn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/JLT.2022.3198649>.

Digital Object Identifier 10.1109/JLT.2022.3198649

For example, the deployment costs of Wi-Fi and Bluetooth technologies are inexpensive, but their positioning accuracy can only achieve the meter level. On the contrary, IR and UWB technologies can achieve centimeter-level positioning accuracy with expensive deployment costs. Considering the drawbacks of indoor positioning technologies mentioned above, moreover, benefiting from the popularity of LED lighting in recent years, visible light positioning has become an emerging and promising technology. Since the LED lighting can be reused with the transmitter of the positioning system, VLP simultaneously meets low deployment cost, high accuracy, energy efficiency, and long lifetime [1]. In addition, a typical VLP system will not generate any radio frequency (RF) signals, which has no interference in RF-sensitive devices and shows good anti-interference for RF signals when in peak time.

Mainstream VLP can be classified into two major categories according to the detector type of receiver, photodiode (PD)-based VLP and image sensor-based VLP. The former is realized with different technologies, including the received signal strength (RSS) (including white light [2] and colors [3]), the angle of arrival (AOA) [4], the time of arrival (TOA) [5], the time difference of arrival (TDOA) [3], [6], the phase difference of arrival (PDOA) [7], and the database-dependent localization [8]. Nevertheless, the RSS based on the optical channel model is susceptible to environmental conditions and the influence of multipath effects cannot be ignored, which degrades the positioning accuracy. In [3], Pergoloni et al. used colored lights to achieve decimeter-level positioning accuracy and compared the performance of schemes based on RSS and TDOA, respectively. Since the angular measurements, the measuring of received signal strength, varying intensities of light, and the solving of quadratic equations jointly determine the result of positioning [9], they usually cause a large error. Besides, high-precision devices are required to estimate the position of the terminal when using the phase difference of arrival (PDOA) or phase of arrival (POA), the time difference of arrival (TDOA) or time of arrival (TOA) [10]. In practical application, three or more LED transmitters are required to realize positioning jointly. Meanwhile, the field of view (FOV) of the PD is smaller than that of the image sensor, and the PD-based VLP system is not suitable for fast moving objects.

Due to the disadvantages mentioned above and the rapid development of high-resolution complementary metal-oxide-semiconductor (CMOS) image sensors, VLP based on CMOS image sensors as an alternative has attracted more and more attention. Fortunately, most commercial smartphones are

equipped with high-resolution CMOS image sensors now, which can be applied in most VLP position schemes. For these VLP systems, multiple lights are captured simultaneously by the image sensor to obtain the geometric features or centers and calculate the location of the receiver. By the rolling shutter effect of CMOS, the on-off state of the LED lights can be captured in the form of stripes and recognized, so that the actual positions of the modulated LEDs can be transmitted. At the same time, the data rate can be raised to a level that is much higher than the capture rate of the image sensor, which is 30 frames per second in general [11]. Moreover, image sensor-based VLP can reach a higher accuracy compared to PD-based positioning schemes because it is less subjected to multipath effects and the diffuse reflection of the light.

A. Motivation

So far, many image sensor-based VLP methods have been proposed, but they commonly have some issues in practical application. For example, in [12], [13], [14], [15], [16], the position methods rely on the assistance of additional device sensors like the gyroscope or IMU to obtain orientations, because the image is needed to combine with the outputs of the additional sensors for positioning or compensation for positioning errors. This is usually a main source of positioning and navigation errors because of the inaccuracy of the azimuth angle shown in previous researches [17], [18]. Then, there is a requirement for the number of transmitters — no less than 3 lights in various schemes [19], [20], [21], [22], [23], [24]. In [25], [26], the authors put forward position methods based on a single LED light with a beacon or a mark. In fact, the beacon or mark here functions as the second light, so they are quasi-double light schemes. If the beacon detection fails, then the positioning will also fail. In addition, the effective positioning area shown in the experiment is too small [26], which is not suitable for large indoor areas with sparse lights. Numerous methods concentrate on the translation between the camera and the anchor light, but few approaches notice the rotation and the intrinsic matrix which is an intrinsic property of the camera in the pinhole camera model. It is usually assumed that the focal length of the camera is already known in many image sensor-based VLP schemes, which plays a critical part in calculating the Z coordinate of the receiver or forming the intrinsic parameter matrix [12], [15], [16], [27], [28], [29], [30], [31]. However, the zoom lens is extensively used to obtain a clear image in most scenarios, thus it's not reasonable to assume the focal length as a constant. Furthermore, some methods only allow a slight tilt when the compensation for angles from an angular sensor [12].

In this paper, we propose a simultaneous localization and calibration VLP method based on double coplanar circular LED lights that can obtain six degrees of freedom (6-DoF), position (x, y, z) and posture (pitch, roll, azimuth), of the camera accurately without additional sensors and calibrate the intrinsic matrix by the focal length. Here, we regarded the focal length as a variable and obtained it by solving an optimization problem of the lamp plane. In the experiment, the ranges of pitch and roll

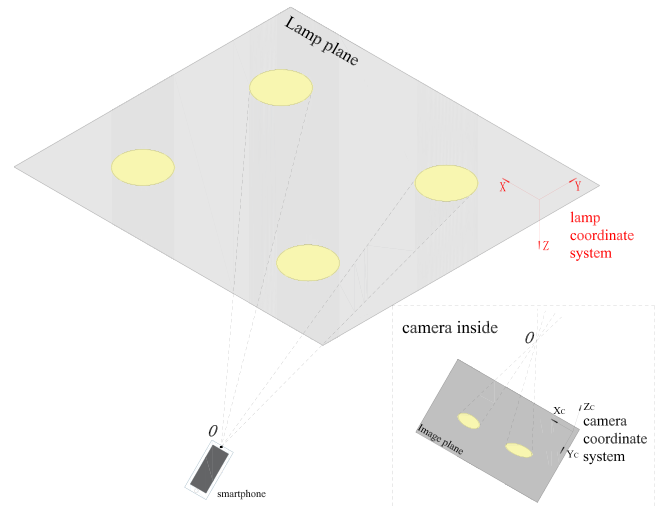


Fig. 1. The schematic diagram of VLP system based on double coplanar LEDs.

angle of the receiver are from about -45° to $+45^\circ$ respectively, which is closer to practical application.

B. Contribution

The main contributions of our research are as follows.

- 1) We relaxed the minimum quantity of LEDs required to two without any beacons or marks and extended the available positioning area of each group of LEDs to 9m^2 .
- 2) The camera receiver can gain 6-DoF in one image and get rid of the dependence on additional sensors.
- 3) The algorithm can dynamically estimate the undetermined focal length, which is applicable in forming the intrinsic matrix when the focal length changes with position.

C. Organization

The rest of the paper is organized as follows. The second section provides a brief description of the system. Then, the third section introduces the model of the positioning algorithm and summarizes the problems to be solved. Next, the fourth section introduces the designed experiment and analyzes the performance of the experiment results quantitatively and qualitatively. Finally, the fifth section draws a conclusion and summarizes the thesis.

II. SYSTEM OVERVIEW

The system of VLP schematically shown in Fig. 1 consists of modulated LED transmitters mounted on the ceiling parallel to the ground and the smartphone receiver. Each LED is assigned a unique identity (ID) that is associated with the actual location. The relations are stored in an ID location database. The coordinate systems of the lamp (denoted by $\{L\}$) and camera (denoted by $\{C\}$) are defined as follows, respectively. The XOY plane of the lamp coordinate system is parallel to the lamp plane, and the Z -axis is defined by the unit normal vector as a reference direction. Similarly, the X_cOY_c plane of the camera coordinate

system is parallel to the image plane, and the Z-axis is defined by the direction of the optical axis, whose corresponding image axis u and v are also parallel to X_c and Y_c , respectively. Two coordinate systems share the same origin which is the optical center of the camera.

The perspective projection of a spatial circle is generally an ellipse. Considering the projection of two coplanar circular LED lights, we can obtain two ellipses on the image sensor plane. When the smartphone and the LED plane are parallel, it is a special case where two ellipses change into two circles. The positioning problem is to estimate the position and the posture of the camera from a perspective image of two ellipses. The posture can be gained from elliptical features and the position can be derived from the translation between the reference lamp and the smartphone. Chen et al. [32] gave a solution to extrinsic parameters in research on camera calibration. In this paper, we extend the method, apply it to VLP, and give a description of the proposed method as follows.

III. PRINCIPLE AND MODEL

In this paper, we simplify the complex nonlinear model of the camera lens system into a simple pinhole camera model in order to analyze the model.

A. The Posture of Camera

The perspective projection of a typical pinhole camera model can be expressed in the below form,

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f/dx & 0 & u_0 & 0 \\ 0 & f/dy & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}, \quad (1)$$

where Z_c represents the z-coordinate of the point with respect to $\{C\}$, $[X_w Y_w Z_w]^T$ is the actual position of the point, and u, v is the coordinate of the point on the image. f (focal length), dx, dy (physical size of the pixels) and u_0, v_0 (the center of the image) constitute the intrinsic matrix, which connects the pixel coordinate system to $\{C\}$. Similarly, T (the translation vector), and R (the rotation matrix) constitute the extrinsic matrix, which connects $\{C\}$ to the world coordinate system. To position the camera with a zoom lens and calibration, the extrinsic matrix and the focal length of the intrinsic matrix need to be figured out.

We consider one circle at first. For any one of the LEDs, it can be expressed in the below form,

$$\begin{cases} (x_l - x_0)^2 + (y_l - y_0)^2 = r^2 \\ z_l = z_0 \end{cases}, \quad (2)$$

where (x_0, y_0) is the center, (x_l, y_l, z_l) are the coordinates of the points on the circle, z_0 is the Z coordinate of points on the plane and the radius is known as r . The equations describe an intersection of an oblique circular cone and the plane $z_l = z_0$. Combining the two expressions above, we can derive the oblique circular cone in the following quadratic form,

$${}^L P_C^T M_C {}^L P_C = 0, \quad (3)$$

where

$$M_C = \begin{bmatrix} 1 & 0 & -\frac{x_0}{z_0} \\ 0 & 1 & -\frac{y_0}{z_0} \\ -\frac{x_0}{z_0} & -\frac{y_0}{z_0} & \frac{x_0^2 + y_0^2 - r^2}{z_0^2} \end{bmatrix},$$

and ${}^L P_C = [x_l \ y_l \ z_l]^T$ is a vector from origin to a point of the circle. In order not to lose generality, we focus on the tilt situation. For a common ellipse in the image, it can be described by the below expression,

$$Au^2 + 2Buv + Cv^2 + 2Du + 2Ev + F = 0. \quad (4)$$

Let the focal length of the camera be denoted by f , by the perspective projection of pinhole camera model,

$$\begin{cases} u = -\frac{fx_c}{z_c} \\ v = -\frac{fy_c}{z_c} \end{cases}.$$

Substituting it into (4), we derive the cone in $\{C\}$,

$$Ax_c^2 + 2Bx_c y_c + Cy_c^2 - \frac{2D}{f} x_c z_c - \frac{2E}{f} y_c z_c + \frac{F}{f^2} z_c^2 = 0,$$

where $[x_c \ y_c \ z_c]^T$ is the coordinates of the points of the ellipse in $\{C\}$, then the above equation can be simplified as,

$${}^C P^T M^C P = 0, \quad (5)$$

where

$$M = \begin{bmatrix} A & B & -\frac{D}{f} \\ B & C & -\frac{E}{f} \\ -\frac{D}{f} & -\frac{E}{f} & \frac{F}{f^2} \end{bmatrix},$$

and ${}^C P = k[x_c \ y_c \ z_c]^T$ is also a vector from origin to a point of the ellipse, $k > 1$ is a scale factor that describes the scale transformation between the two vectors. Since M is a real symmetric matrix, it can be diagonalized by an orthonormal matrix U subjected to $U^T U = I$,

$$M = U \Lambda U^T, \quad (6)$$

where $\Lambda = \text{Diag}\{m_1, m_2, m_3\}$ are the eigen-values and $U = [u_1, u_2, u_3]$ are the normalized eigen-vectors. Since $\{L\}$ and $\{C\}$ have the same origin, there exists a rotation matrix ${}^C_L R$ that transforms ${}^L P_C$ to ${}^C P$ as following,

$${}^C P = {}^C_L R {}^L P_C. \quad (7)$$

Substituting (6) and (7) into (5) and combining (3) can derive the following relation,

$$k {}^C_L R^T U \Lambda U^T {}^C_L R = M_C. \quad (8)$$

We define $R_U = U^T {}^C_L R$, which is also a rotation matrix. So, we have

$$k R_U^T \Lambda R_U = M_C, \quad (9)$$

where $R_U = \begin{bmatrix} r_{1x} & r_{2x} & r_{3x} \\ r_{1y} & r_{2y} & r_{3y} \\ r_{1z} & r_{2z} & r_{3z} \end{bmatrix}$. Noticing the $I_{2 \times 2}$ in the upper left corner of M_C , we have constraints from (9),

$$\begin{cases} m_1 r_{1x}^2 + m_2 r_{1y}^2 + m_3 r_{1z}^2 = m_1 r_{2x}^2 + m_2 r_{2y}^2 + m_3 r_{2z}^2 \\ m_1 r_{1x} r_{2x} + m_2 r_{1y} r_{2y} + m_3 r_{1z} r_{2z} = 0 \end{cases}. \quad (10)$$

When using the similarity and substituting (6) into (5), another constraint of the cone is that,

$$m_1 x_u^2 + m_2 y_u^2 + m_3 z_u^2 = 0,$$

where $U^T C P = [x_u, y_u, z_u]^T$, compared with the standard equation of the quadratic cone, without loss of generality,

$$\begin{cases} m_1 m_2 > 0 \\ m_2 m_3 < 0 \\ |m_1| \geq |m_2| > 0 \end{cases}. \quad (11)$$

Considering (10) and unitizing the vectors of R_U , R_U also has the constraints as follows,

$$\begin{cases} r_{1x}^2 + r_{1y}^2 + r_{1z}^2 = 1 \\ r_{2x}^2 + r_{2y}^2 + r_{2z}^2 = 1 \\ r_{3x}^2 + r_{3y}^2 + r_{3z}^2 = 1 \end{cases}, \quad (12)$$

so the parameters corresponding to the circle on the right side of (9) can be derived,

$$\begin{cases} \frac{x_0}{z_0} = -k(m_1 r_{1x} r_{3x} + m_2 r_{1y} r_{3y} + m_3 r_{1z} r_{3z}) \\ \frac{y_0}{z_0} = -k(m_1 r_{2x} r_{3x} + m_2 r_{2y} r_{3y} + m_3 r_{2z} r_{3z}) \\ \frac{x_0^2 + y_0^2 - r^2}{z_0^2} = k(m_1 r_{1x} r_{2x} + m_2 r_{1y} r_{2y} + m_3 r_{1z} r_{2z}) \end{cases}. \quad (13)$$

By simplifying (13), we can get the following results,

$$\begin{cases} k = m_2 \\ \frac{x_0}{z_0} = -t_2 \frac{\sqrt{(m_1 - m_2)(m_2 - m_3)} \cos \theta}{m_2} \\ \frac{y_0}{z_0} = -t_1 t_2 \frac{\sqrt{(m_1 - m_2)(m_2 - m_3)} \sin \theta}{m_2} \\ z_0^2 = -\frac{r^2 m_2^2}{m_1 m_3} \end{cases}, \quad (14)$$

where θ is an auxiliary variable without practical significance, with undetermined signs t_1, t_2 . The unit normal vector of the light plane in $\{C\}$, denoted by \mathbf{N} , is parallel to the Z-axis of $\{L\}$ and the center of the light in $\{C\}$ is denoted by \mathbf{C} , so we can derive them from what they are in $\{L\}$,

$$\begin{cases} \mathbf{C} = {}^C_L R \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = U R_U \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = z_0 U \begin{bmatrix} t_2 \frac{m_3}{m_2} L \\ 0 \\ -t_1 \frac{m_1}{m_2} J \end{bmatrix} \\ \mathbf{N} = {}^C_L R \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = U R_U \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = U \begin{bmatrix} t_2 L \\ 0 \\ -t_1 J \end{bmatrix} \end{cases}, \quad (15)$$

where $z_0 = \frac{r m_2}{\sqrt{-m_1 m_3}}$ and $\begin{cases} J = \sqrt{\frac{m_2 - m_3}{m_1 - m_3}} \\ L = \sqrt{\frac{m_1 - m_2}{m_1 - m_3}} \end{cases}$.

To determine the signs in (15), we denote the direction coming out of the light plane as the positive direction of the normal vector \mathbf{N} , and the circle faces to the camera. Then, we have the following inequality constraints,

$$\begin{cases} \mathbf{C} \cdot [0 \ 0 \ 1]^T > 0 \\ \mathbf{N} \cdot [0 \ 0 \ 1]^T < 0 \end{cases}. \quad (16)$$

So, there is no more than two sets of possible solutions of \mathbf{N} and \mathbf{C} , and only one solution whose \mathbf{C} in front of the camera is acceptable in our research. To figure out the \mathbf{N} is to find a plane whose intersection with the above conical surface is a circle. Since the focal length f is unknown, it is retained in all expressions. Because the two LEDs are coplanar, the unit normal

vectors of the LEDs, denoted by $N_1(f)$ and $N_2(f)$, would be parallel,

$$N_1(f) = N_2(f).$$

Therefore, the solution to f can be transformed into the following optimization problem,

$$\min_f y = (N_1(f) \cdot N_2(f) - 1)^2, \quad (17)$$

from which we can get the f and figure out the \mathbf{N} and \mathbf{C} .

Here, we select the center \mathbf{C}_1 of reference light as the origin to establish the positioning coordinate system, denoted by $\{P\}$, which can be defined by the vectors in $\{C\}$ as follows,

$$\begin{cases} i = \frac{\mathbf{C}_2 - \mathbf{C}_1}{|\mathbf{C}_2 - \mathbf{C}_1|} \\ k = N \\ j = k \times i \end{cases}, \quad (18)$$

where i, j, k are the bases of $\{P\}$, and \mathbf{C}_2 is the center of the other light. Thus, the orientation of $\{P\}$ described in $\{C\}$ can be given by

$$\begin{aligned} {}^C_P R_W &= [i \ j \ k] \\ &= \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \\ &\times \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}, \end{aligned} \quad (19)$$

where α, β, γ represent pitch, roll and azimuth angle, respectively.

B. The Position of Camera

If we know a vector ${}^P P_A$ described with respect to $\{P\}$, it is easy to compute it with respect to $\{C\}$ by the following expression,

$$\begin{bmatrix} {}^C P_A \\ 1 \end{bmatrix} = \begin{bmatrix} {}^C_P R & {}^C T_{PORG} \\ 0_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} {}^P P_A \\ 1 \end{bmatrix}, \quad (20)$$

where ${}^C T_{PORG}$ locates the origin of $\{P\}$ relative to $\{C\}$. Hence, the position of camera with respect to $\{P\}$ can be computed by

$${}^P P_{camera} = -{}^P_C R_W {}^C P_{C1} = -{}^C_P R_W^T {}^C P_{C1}, \quad (21)$$

where ${}^C P_{C1}$ is the coordinate of reference light \mathbf{C}_1 with respect to $\{C\}$. Then the actual position of the camera can be obtained by referring to the actual position of the reference light.

C. The Focal Length and Intrinsic Matrix

To solve the optimization problem in (17), a sensible approach is to find the range of solution. Before positioning, we estimate the range of focal length firstly. If the camera is parallel to the lamp plane, by the similarity property of triangles, we have

$$\frac{D}{H} = \frac{d}{f_0}, \quad (22)$$

where D is the distance between two centers of lights, H is the height between the optical center and the lights, d is the distance between two centers of circles on the image, and f_0 is the approximate focal length. Hence, we can obtain the solution range $[(1-t)f_0, (1+t)f_0]$ by a scale t , $0 < t \leq 1$, and the

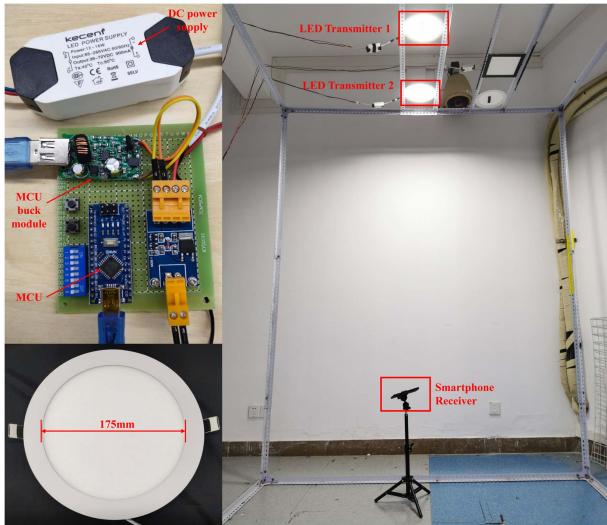


Fig. 2. The environment and equipment of static experiment.

solution to focal length f in (17) can be computed by iteration. The other intrinsic parameters in (1) that are constants can be measured in advance, so we can form the intrinsic matrix by this solution and achieve the calibration.

IV. EXPERIMENT AND ANALYSIS

A. Static Experiment Settings

We evaluated the accuracy and effectiveness of the proposed method through experiments in the following. The whole experiment area is $3\text{m} \times 3\text{m}$, divided into $0.5\text{m} \times 0.5\text{m}$ grid size, and 49 test points are selected at the height of approximately 2m. The LED transmitter consists of the control module and the LED light. As is shown in Fig. 2, two coplanar circular LEDs are mounted parallel to the ground and the distance between the two centers of LEDs is 600mm. In order to be scalable, the control circuit unit is made up of several off-the-shelf modules. The AC to DC power supply converts AC 220V to DC 72V, providing a power supply for the LED driver module and the Microcontroller Unit (MCU) buck module. Then, the MCU driven by the MCU buck module generates the control signal to control the amplification module to drive the LED transmitter. Next, the ID information of each light is downloaded into the memory of the MCU in advance, converted into high and low levels, and generates the control signal to modulate the LEDs by adopting on-off keying (OOK) modulation scheme, which achieves the signal transmission.

The receiver is an IQOO smartphone with a 12-Megapixel front camera. Before positioning, we calibrated the intrinsic parameters of the camera except focal length. Then, we set the camera horizontal to estimate the range of focal length mentioned in Section III C. Here, we took $t = 0.5$ initially. Given that the actual experimental environment is not ideal, we considered that the focal length is effective if $i \cdot k < 0.2$ in (19). To avoid accidental errors, if we cannot find a satisfactory result in this interval that $t = 0.5$ for five consecutive times, we increase the value of t , $t = t + 0.1$, until an effective solution

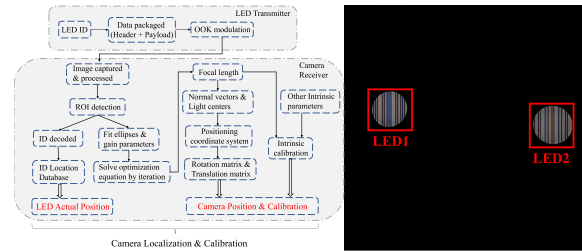


Fig. 3. The diagram of the position system and the captured image of LEDs.

can be found. If an effective solution still cannot be obtained when $t = 1$, it indicates that the camera needs to be recalibrated at another height to estimate the approximate focal length and the solution interval. To reduce the effect of light saturation, we minimized the ISO and the exposure time of the camera. By the rolling shutter effect, the transmitted signal is obtained in the form of light and dark image stripes. We used Gaussian blur and erosion operations to eliminate the blooming and smear on the image. Next, the image is processed by RGB to grayscale, binarization, and closed operation to eliminate the influence of stripes, so we gained ROI of the LED images by LED-ID detected method from the source image, based on which the ID is identified by an efficient decoding scheme. Hence, referring to the established ID location database that is stored in the receiver, different LEDs can be easily distinguished and the actual locations of the LEDs can also be obtained. Meanwhile, after image processing, we used the least square method to fit the contours of LEDs extracted in ROI and gained the parameters of two ellipses that are input to the positioning algorithm in Section III. We solved it by simulated annealing method. After solving the optimization problem, we can get the focal length to calculate the LED centers and the normal vectors of the lamp plane, and then the rotation matrix and translation vector are obtained. Thus, we can obtain the relative position of the receiver with respect to the reference LED (specified according to the ID), and combining the actual location of the reference LED mentioned above can finally obtain the actual position of the receiver. If the estimated ID did not match the ID in the database, the positioning result will not update. In this case, when limiting the usage of ID codes, there is only a small probability that an incorrect ID matches another correct ID in the database. By this means, we can ensure the accuracy of positioning. When multiple lights appear in the image, we choose the two LEDs that are nearest to us to position. If we are at the same distance from the two lights, we choose one of them as the reference LED according to the numerical value of the ID code. The algorithm diagram of the positioning system and the captured image of LEDs can be seen in Fig. 3. The specific processes of the efficient decoding scheme are mentioned in [33], and the details of the equipment are shown in Table I.

B. Static Performance

In this section, we focus on the static performance of our positioning algorithm. In order to prove the effectiveness and

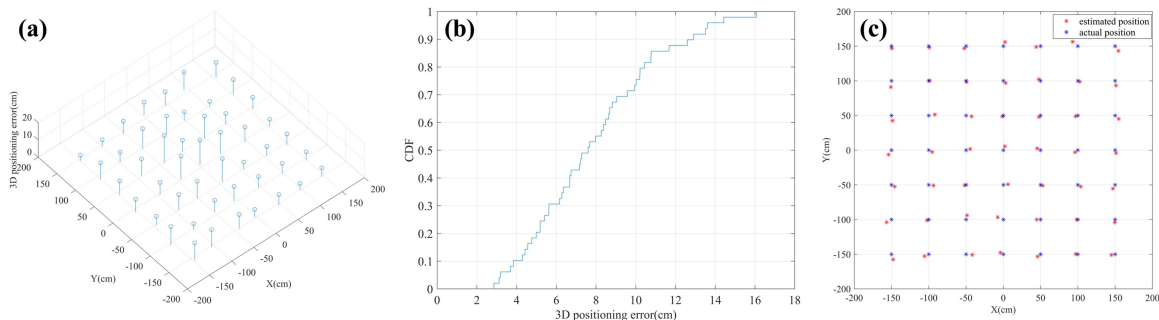


Fig. 4. The errors of position results. (a) The 3D absolute errors of each position point. (b) The cumulative density function (CDF) of 3D absolute errors. (c) The horizontal position results (red star) ignoring the Z-axis direction and the actual point (blue star).

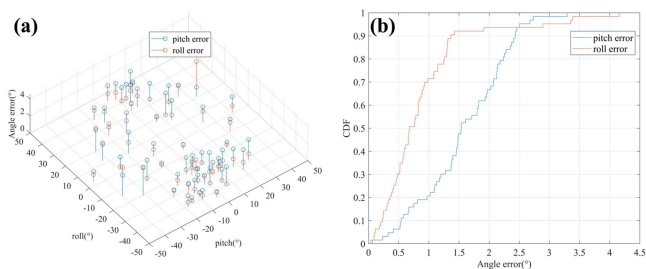


Fig. 5. (a) The mean angle errors of the same pitch and roll at different points and (b) the corresponding cumulative density function (CDF).

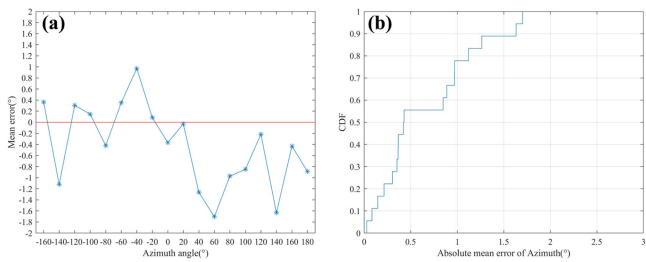


Fig. 6. The azimuth mean errors (a) and the CDF (b) (pitch = 10.0°, roll = 10.0°).



Fig. 7. The experiment environment and the map of indoor dynamic experiment.

give a quantitative evaluation of our method, we analyzed the accuracy of the position and posture results.

1) *Position Accuracy*: To avoid accidental errors, we adopted the average positioning results calculated from ten images with three different angles at each point as position results, based on which we calculated the positioning errors. The corresponding

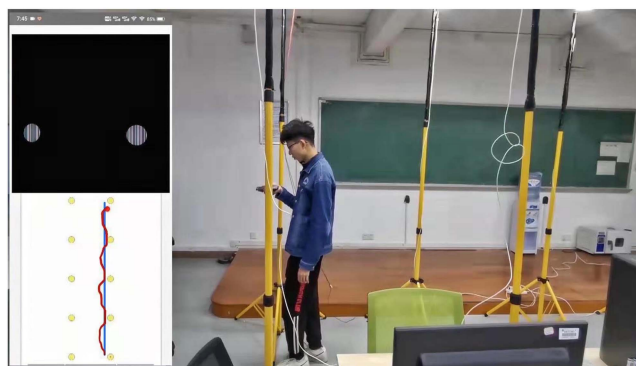


Fig. 8. The positioning track (red line) and the actual track (blue line) of the experiment.

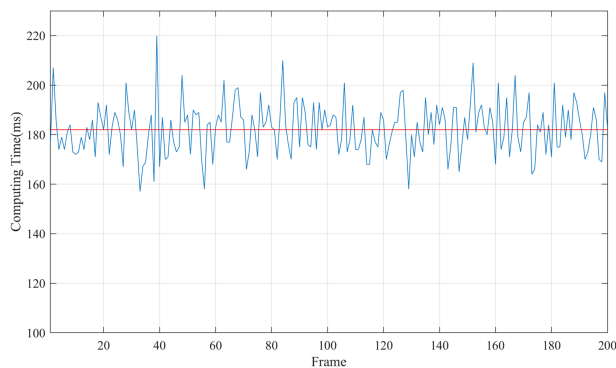


Fig. 9. The computing time and the average (red line) of our VLP method.

TABLE I
PARAMETERS OF VLP SYSTEM

Transmitter specifications	
Working Voltage of LED/V	36-72 DC
Power of LED/w	18
Diameter of LED/mm	175
MCU	ATMELATMEGA328P-AU
Coordinate of LED1/mm	(0, 0, 2500)
Coordinate of LED2/mm	(600, 0, 2500)
Receiver specifications	
Module	IQOO front camera
Processor	Qualcomm® Snapdragon 855
Resolution	4032×3024, 12 Megapixel
Format of image	JPG
Time of Exposure	1/12000s

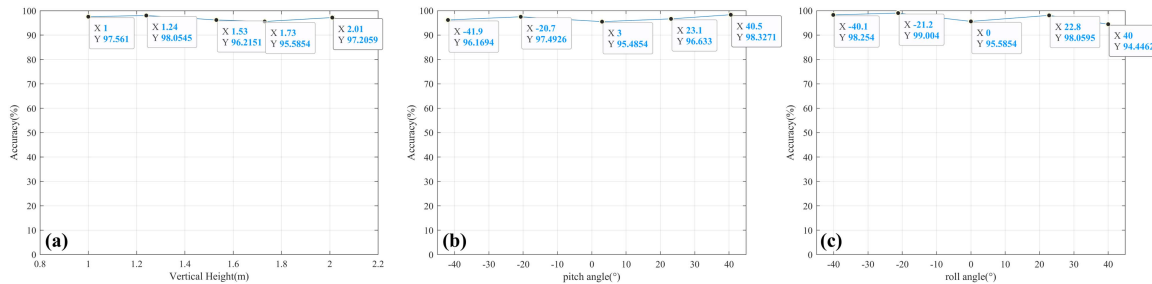


Fig. 10. The accuracy of ID decoded. (a) The accuracy of different vertical height (pitch=0°, roll=0°). (b) The accuracy of different pitch angles (vertical height=1.75m). (c) The accuracy of different roll angles (vertical height=1.75m).

TABLE II
ERRORS OF ORIENTATIONS

Orientation	Mean (°)	Min (°)	Max (°)	Standard deviation (°)
Pitch	1.58	0.04	3.29	0.70
Roll	0.89	0.07	4.16	0.78
Azimuth	0.67	0.03	1.70	0.52

absolute errors for each point are shown in Fig. 4(a). The error cumulative density function is shown in Fig. 4(b). These data suggest that the 3D positioning error ranges from 2.86cm to 16.08cm, with a mean error of 7.91cm and a standard deviation of 3.16cm when the vertical height is about 2m. Generally, we noticed that the largest 3D positioning errors occurred near the center of the experiment area, which mainly concentrated on the Z-axis direction. It is because the ellipses tend to circles when the receiver is near the center and the orientation also tends to be horizontal, the solution to the lamp plane becomes divergent, resulting in the decline of precision of the normal vector. Hence, when the errors in the Z-axis direction were ignored, the horizontal positioning results are shown in Fig. 4(c). Overall, the positions estimated by this method matched well with the actual positions.

2) *Posture Accuracy*: The errors and the cumulative density function (CDF) of each orientation (pitch, roll, and azimuth) are shown in Figs. 5 and 6. Fig. 5. shows the mean angle errors of the same pitch and roll at different points from about -45° to +45°, which is close to reality. The angle error is not significantly related to the value itself. The CDF suggests that most of the results are close to the real value, except for a few outliers. Differently from the former, the mean azimuth errors were calculated from 10 points at different positions in the same orientation. The error of the azimuth angle is relatively small, which is often the main source of positioning error in the methods relying on the angular sensors. The specific numerical description of errors is shown in Table II.

3) *Focal Length Effectiveness*: In addition, the estimated focal length is also obtained by solving the problem. Since it is impossible to measure the focal length of a smartphone directly, we are not able to give a quantitative evaluation of the error of focal length. However, as an alternative, we can evaluate its effectiveness indirectly in the experiments. The focal length is used to calculate the normal vectors of the lamp plane, and then to find the rotation matrix and the translation

vector. When the estimated focal length deviates from the actual value, the planes of two lamps reconstructed by it will not be completely coplanar. If the error is too large, then the positioning will be ineffective. Therefore, we can use the bases in (19) to describe how coplanar the planes of two lamps are, so as to evaluate how the focal length deviates from the actual value. If we derive an incorrect focal length, the bases of the rotation matrix will not be orthogonal. We considered that the focal length is effective if $i \cdot k < 0.2$ in the rotation matrix. In this case, the error of the included angle between i and k can be up to 11.5 degrees, but we still got high accuracy position and posture results above, which showed that when the error of focal length within a limited range will not significantly affect the positioning accuracy and proved the robustness and effectiveness. Hence, we considered that the intrinsic matrix calibrated here is acceptable, too.

C. Indoor Dynamic Experiment Settings

To further verify the dynamic effectiveness and real-time performance of the algorithm, ten LED transmitters mounted on the flat plates are applied to simulate a real application environment. The details of the experiment environment can be seen in Fig. 7, and the devices are the same as the static experiment mentioned above. In the experiment, we made the two lights in the same row as a group to position the walker with a smartphone. Then we walked through the illumination area in a straight line and recorded the track of the position results. The actual track of the walker is plotted as a blue line and the track of the position results is plotted as a red line. The demonstration video of our proposed VLP system is available on our website.¹

D. Dynamic Performance

In this section, we compared the results of the positioning track with the actual track and discussed the real-time performance of the system.

The method is an online method, so we can monitor the output of positioning in real-time. As we can see in Fig. 8, the positioning error increases as the position away from the reference light and the tilt increases. When switching to the next group of lights, the offset was corrected and the error was eliminated,

1.[Online]. Available: https://www.bilibili.com/video/BV1wb4y1E7Kd?spm_id_from=333.999.0.0

TABLE III
PERFORMANCE COMPARISON WITH THE SOTA

Method	Average Accuracy (cm)	Time Consumption (ms)	Require LEDs (at least)	Hardware	Degree of Freedom	Receiver Type	Density of LEDs
R. [34]	3.9	44	2	Raspberry+computer	3D+yaw	Camera	9 -(1.8×1.8×2.0)
R. [10]	0.86	162	1	computer	2D	Camera	4 -(1.9×1.0×1.9)
R. [12]	<6.5	<i>N/A</i>	2	Smartphone	2D	Camera + gyroscope	2 -(0.9×1.1×2.45)
R. [13]	7.3&67.2	<i>N/A</i>	1	Smartphone	3D	Camera + gyroscope	1 -(1.0×1.0×3.0)
R. [16]	16 (90%)	<i>N/A</i>	1	Smartphone	6D	Camera+IMU	1 -(1.8×1.8×2.0)
R. [20]	1	<i>N/A</i>	3	Smartphone+computer	2D	Camera	3 -(0.75×1.05×2.3)
R. [25]	17.5	<i>N/A</i>	1*	Smartphone	2D+3D	Camera	1 -(3.0×3.0×2.0)
R. [28]	13.4	<i>N/A</i>	1	Smartphone	3D+yaw	Camera + PDR	1 -(1.0×1.0×2.6)
R. [27]	2.2- 3.0	72.4-152.1 [#]	1	Raspberry+computer	2D+yaw	Camera+gyroscope+encoder	4 -(2.0×2.0×1.5)
R. [26]	2.26	60	1*	Ingenic X1500 (CPU)	2D	Camera	1 -(1.6×1.6×3.0)
R. [23]	6.58	<i>N/A</i>	4	Smartphone	3D	Camera	4 -(1.0×1.0×0.8)
R. [24]	6.6	<i>N/A</i>	3	Camera	2D	Camera	3 -(0.9×0.3×1.8)
R. [29]	2.1	33	1 or 0	Raspberry+computer	3D+yaw	SLOVLP + IMU	3 -(9.1×4.0×2.7)
R. [31]	3.85	56.28	2	computer	2D+yaw	Camera	2 -(0.8×0.8×2.0)
R. [15]	5.5	<i>N/A</i>	1	Smartphone	3D	Camera + IMU	1 -(2.7×1.8×1.5)
ours	7.91	182	2	Smartphone	6D	Camera	2 -(3.0×3.0×2.0)

* means that the LED lamp with beacon marker.

means that the position estimation was performed in post-processing, not real-time.

which prevented positioning errors from accumulating. In this demonstration, the positioning track is roughly close to the actual track.

Real-time performance is also a key factor for positioning systems, especially for low-power mobile devices. On the one hand, the positioning results need to respond quickly to the changes in the actual position of the receiver. On the other hand, the performance of low-power mobile devices is insufficient, so a too complex algorithm will consume a large amount of computing time, resulting in high latency. In our experiment, we continuously measured the computing time of our method 200 times to calculate the average time consumption, which is shown in Fig. 9. The computing time ranges from 157ms to 220ms, with an average of 182ms. To reduce the burden of the smartphone, we updated the position result every 500ms in the demonstration.

It is worth mentioning that the changes in angle and height will not affect the decoding accuracy, which ensures the good running of the VLP system. Only two lights of a group decoded correctly can keep the system running correctly. As is shown in Fig. 10(a), when the height ranges from 1m to 2m, the decoding accuracy continuously measured over 500 times at each height was not less than 95%. Also, in Fig. 10(b) and 10(c), the decoding accuracy measured over 500 times at each angle was almost above 95%, when the pitch and roll ranged from about -40° to 40° and the vertical height was 1.75m. Therefore, our proposed method also has good stability in most cases.

E. Discussion

Here, we compared the performance of our simultaneous localization and calibration VLP method with the state-of-the-art (SOTA) works in the field in Table III. Our method gave the 6-DOF of the receiver and intrinsic parameter calibration without the assistance of additional sensors, which was not achieved by other works in Table III. Compared with another 6-DOF positioning method [16], our approach achieved higher position and posture accuracy, especially the high-precision azimuth angle, and the density of LEDs per unit area was lower.

Actually, our density was only higher than a few methods [15], [25], [29]. Furthermore, we uniquely gave the computing time that covers both the ID decoding and positioning processes on the smartphone receiver to show the real-time performance. And we carried out the dynamic experiment to demonstrate the actual performance. In [27], [34], the authors also gave comparisons of the positioning trajectory and actual trajectory, but the positioning area of our experiment was larger. Meanwhile, our study that only requires a camera as a receiver and several modulated LED transmitters to achieve precision positioning reduces the deployment cost and the expense of precision additional sensors of VLP, which makes it possible for large-scale application. In the demonstration, the instability of the receiver brought by moving will not affect the accuracy greatly. In summary, the delivered performance satisfies the needs of a number of practical applications.

V. CONCLUSION

In this paper, we proposed a simultaneous localization and calibration VLP method based on double coplanar circular LED lights, and discuss the accuracy of position and posture and real-time performance. We made full use of the geometric features of two LED projections to get rid of the dependence on additional sensors and solved localization and calibration problems. Our research will help extend VLP applied to mobile devices without angular sensors, reducing the cost and threshold of the technology application. The experiment result showed that our method had high 3D accuracy (7.91cm), high angle accuracy (<1.6°), and low latency (182ms), which can satisfy the requirements of mobile positioning. In the future, we will combine the inertial measurement unit to further extend the positioning area and improve the positioning accuracy by Extended Kalman Filter.

REFERENCES

- [1] J. Armstrong, Y. A. Sekercioglu, and A. Neild, "Visible light positioning: A roadmap for international standardization," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 68–73, Dec. 2013, doi: [10.1109/MCOM.2013.6685759](https://doi.org/10.1109/MCOM.2013.6685759).

- [2] M. Angelichinoski, D. Denkovski, V. Atanasovski, and L. Gavrilovska, "Cramér–Rao lower bounds of RSS-based localization with anchor position uncertainty," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2807–2834, May 2015.
- [3] S. Pergoloni, Z. Mohamadi, A. M. Vegni, Z. Ghassemlooy, and M. Biagi, "Metameric indoor localization schemes using visible lights," *IEEE, J. Lightw. Technol.*, vol. 35, no. 14, pp. 2933–2942, Jul. 2017, doi: [10.1109/JLT.2017.2706527](https://doi.org/10.1109/JLT.2017.2706527).
- [4] K. Dogançay and H. Hmam, "Optimal angular sensor separation for AOA localization," *Signal Process.*, vol. 88, pp. 1248–1260, May 2008.
- [5] I. Guvenc and C.-C. Chong, "A survey on TOA based wireless localization and NLOS mitigation techniques," *IEEE Commun. Surv. Tut.*, vol. 11, no. 3, pp. 107–124, Jul.–Sep. 2009.
- [6] T.-H. Do and M. Yoo, "TDOA-based indoor positioning using visible light," *Photon. Netw. Commun.*, vol. 27, no. 2, pp. 80–88, 2014.
- [7] A. Naz, H. M. Asif, T. Umer, and B.-S. Kim, "PDOA based indoor positioning using visible light communication," *IEEE Access*, vol. 6, pp. 7557–7564, 2018.
- [8] M. Biagi, A. M. Vegni, and T. D. C. Little, "LAT indoor MIMO-VLC — Localize, access and transmit," in *Proc. IEEE, Int. Workshop Opt. Wireless Commun.*, 2012, pp. 1–3, doi: [10.1109/IWOW.2012.6349698](https://doi.org/10.1109/IWOW.2012.6349698).
- [9] W. Guan, S. Wen, L. Liu, and H. Zhang, "High-precision indoor positioning algorithm based on visible light communication using complementary metal–oxide–semiconductor image sensor," *Opt. Eng.*, vol. 58, 2019, Art. no. 024101.
- [10] W. Guan, X. Chen, M. Huang, Z. Liu, Y. Wu, and Y. Chen, "High-speed robust dynamic positioning and tracking method based on visual visible light communication using optical flow detection and Bayesian forecast," *IEEE Photon. J.*, vol. 10, no. 3, Jun. 2018, Art. no. 7904722.
- [11] 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, 2018.
- [12] J.-Y. Kim, S.-H. Yang, Y.-H. Son, and S.-K. Han, "High-resolution indoor positioning using light emitting diode visible light and camera image sensor," *IET Optoelectron.*, vol. 10, no. 5 pp. 184–192, 2016.
- [13] Y. Hou, S. Xiao, M. Bi, Y. Xue, W. Pan, and W. Hu, "Single LED beacon-based 3-D indoor positioning using off-the-shelf devices," *IEEE Photon. J.*, vol. 8, no. 6, Dec. 2016, Art. no. 6806211.
- [14] H. Huang, L. Feng, G. Ni, and A. Yang, "Indoor imaging visible light positioning with sampled sparse light source and mobile device," *Chin. Opt. Lett.*, vol. 14, 2016, Art. no. 090602.
- [15] H. Cheng, C. Xiao, Y. Ji, J. Ni, and T. Wang, "A single LED visible light positioning system based on geometric features and CMOS camera," *IEEE Photon. Technol. Lett.*, vol. 32, no. 17, pp. 1097–1100, Sep. 2020, doi: [10.1109/LPT.2020.3012476](https://doi.org/10.1109/LPT.2020.3012476).
- [16] J. Hao, J. Chen, and R. Wang, "Visible light positioning using a single LED luminaire," *IEEE Photon. J.*, vol. 11, no. 5, Oct. 2019, Art. no. 7905113, doi: [10.1109/JPHOT.2019.2930209](https://doi.org/10.1109/JPHOT.2019.2930209).
- [17] U. Shala and A. Rodriguez, "Indoor positioning using sensor-fusion in android devices," M.S. thesis, Dept. Comput. Sci., Kristianstad Univ., Kristianstad, Sweden, 2011.
- [18] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao, "A reliable and accurate indoor localization method using phone inertial sensors," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 421–430.
- [19] M. Yoshino, S. Haruyama, and M. Nakagawa, "High-accuracy positioning system using visible LED lights and image sensor," in *Proc. IEEE Radio Wireless Symp.*, 2008, pp. 439–442, doi: [10.1109/RWS.2008.4463523](https://doi.org/10.1109/RWS.2008.4463523).
- [20] J. Xu, C. Gong, and Z. Xu, "Experimental indoor visible light positioning systems with centimeter accuracy based on a commercial smartphone camera," *IEEE Photon. J.*, vol. 10, no. 6, Dec. 2018, Art. no. 7908717, doi: [10.1109/JPHOT.2018.2878532](https://doi.org/10.1109/JPHOT.2018.2878532).
- [21] M. S. Rahman, M. M. Haque, and K. D. Kim, "Indoor positioning by LED visible light communication and image sensors," *Int. J. Elect. Comput. Eng.*, vol. 1, pp. 420–431, 2011.
- [22] Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, and P. L. Dutta, "Indoor positioning with mobile phones and visible light," in *Proc. Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 447–458.
- [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] B. Lin, Z. Ghassemlooy, C. Lin, X. Tang, Y. Li, and S. Zhang, "An indoor visible light positioning system based on optical camera communications," *IEEE Photon. Technol. Lett.*, vol. 29, no. 7, pp. 579–582, Apr. 2017.
- [25] R. Zhang, W. Zhong, Q. Kemao, and S. Zhang, "A single LED positioning system based on circle projection," *IEEE Photon. J.*, vol. 9, no. 4, Aug. 2017, Art. no. 7905209.
- [26] H. Li et al., "A fast and high-accuracy real-time visible light positioning system based on single LED lamp with a beacon," *IEEE Photon. J.*, vol. 12, no. 6, Dec. 2020, Art. no. 7906512, doi: [10.1109/JPHOT.2020.3032448](https://doi.org/10.1109/JPHOT.2020.3032448).
- [27] R. Amsters, D. Holm, J. Joly, E. Demeester, N. Stevens, and P. Slaets, "Visible light positioning using Bayesian filters," *J. Lightw. Technol.*, vol. 38, no. 21, pp. 5925–5936, Nov. 2020, doi: [10.1109/JLT.2020.3006618](https://doi.org/10.1109/JLT.2020.3006618).
- [28] H. Huang et al., "Hybrid indoor localization scheme with image sensor-based visible light positioning and pedestrian dead reckoning," *Appl. Opt.*, vol. 58, no. 12, pp. 3214–3221, 2019.
- [29] W. Guan, L. Huang, B. Hussain, and C. P. Yue, "Robust robotic localization using visible light positioning and inertial fusion," *IEEE Sensors J.*, vol. 22, no. 6, pp. 4882–4892, Mar. 2022, doi: [10.1109/JSEN.2021.3053342](https://doi.org/10.1109/JSEN.2021.3053342).
- [30] S. Chen and W. Guan, "High accuracy VLP based on image sensor using error calibration method," 2020. *arXiv:2010.00529*.
- [31] W. Guan, X. Zhang, Y. Wu, Z. Xie, J. Li, and J. Zheng, "High precision indoor visible light positioning algorithm based on double LEDs using CMOS image sensor," *Appl. Sci.*, vol. 9, no. 6, 2019, Art. no. 1238. [Online]. Available: <https://doi.org/10.3390/app9061238>
- [32] Q. Chen, H. Wu, and T. Wada, "Camera calibration with two arbitrary coplanar circles," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 521–532.
- [33] H. Song, S. Wen, C. Yang, D. Yuan, and W. Guan, "Universal and effective decoding scheme for visible light positioning based on optical camera communication," *Electronics*, vol. 10, no. 16, 2021, Art. no. 1925. [Online]. Available: <https://doi.org/10.3390/electronics10161925>
- [34] P. Lin et al., "Real-time visible light positioning supporting fast moving speed," *Opt. Exp.*, vol. 28, no. 10, pp. 14503–14510, 2020.
- [35] W. Guan et al., "Robot localization and navigation using visible light positioning and SLAM fusion," *J. Lightw. Technol.*, vol. 39, no. 22, pp. 7040–7051, Nov. 2021, doi: [10.1109/JLT.2021.3113358](https://doi.org/10.1109/JLT.2021.3113358).
- [36] W. Guan, S. Chen, S. Wen, Z. Tan, H. Song, and W. Hou, "High-accuracy robot indoor localization scheme based on robot operating system using visible light positioning," *IEEE Photon. J.*, vol. 12, no. 2, Apr. 2020, Art. no. 7901716, doi: [10.1109/JPHOT.2020.2981485](https://doi.org/10.1109/JPHOT.2020.2981485).

Junye Chen is currently working toward the B.E. degree with the South China University of Technology, Guangzhou, China (will receive in 2023), major in intelligence science and technology. His research interests include visible light positioning and cooperative localization.

Dinghao Zeng received the B.E. degree from the School of Automation Science and Engineering, South China University of Technology, Guangzhou, China, in 2022. His research interests include picture processing and visible light positioning algorithm.

Chen Yang is currently working toward the B. E. degree from the South China University of Technology, Guangzhou, China (will receive in 2023), major in intelligence science and engineering. Her research interests include picture processing and visible light positioning algorithm.

Weipeng Guan received the B.E. degree in electronic science and technology from the School of Materials Science and Engineering, and the M.E. degree in control theory and control engineering from the School of Automation Science and Engineering, South China University of Technology, China, in 2016 and 2019, respectively. He is currently working toward the Ph.D. degree. He was with Samsung Electronics Company Ltd., Huawei Technologies Company Ltd., TP-LINK, the Chinese Academy of Sciences, and The Chinese University of Hong Kong. He has authored or coauthored more than 50 research articles in prestigious international journals and conferences. His research interests include mobile robot, optical camera communications, visible light positioning, and sensor fusion.