A VLC Smartphone Camera based Indoor Positioning System

Yiwei Li, Zabih Ghassemlooy, *Senior Member*, *IEEE*, Xuan Tang, Bangjiang Lin and Yi Zhang

***Abstract*—We present a real-time indoor visible light positioning system based on optical camera communication, where the coordinate data in the on-and-off keying (OOK) format is transmitted via light emitting diode (LED) based lights and captured using a smartphone camera. The position of camera is estimated using a novel Perspective-*n*-Point (P*n*P) problem algorithm, which determines the position of a calibrated camera from *n* 3D-to-2D point correspondences. The experimental results show that the proposed system offers mean position errors of 4.81 cm and 6.58 cm for the heights of 50 cm and 80 cm, respectively.**

***Index Terms*—Indoor positioning systems, perspective-*n*-point algorithm, optical camera communication.**

I. Introduction

With the growing research in the field of indoor positioning, we have seen the introduction of a new type of indoor positioning system (IPS) based on visible light communications (VLC) [1]. VLC uses the LED based lighting fixtures to provide high positioning accuracy, which is low cost, license-free and not susceptible to the radio frequency (RF) electromagnetic interference [2]. Additionally, in VLC based IPS, the link is less subjected to the multipath effects compared to the RF based systems, thus making propagation of light beams more predictable [3]. More importantly, with the indoor lighting infrastructures being ubiquitous, the realization of VLC-IPS system has becomes feasible with a high level of positioning accuracy.

This letter classifies the state-of-the-art VLC based IPS into five groups: proximity, fingerprinting, dead-reckoning, triangulation and vision analysis. The proximity and fingerprinting techniques, which are simple to implement and with a low positioning accuracy compared to the other schemes. Note that, the proximity method only provides proximity location information based on the signal from a single LED base station with a unique identification code [4]. The positioning error (PE) in this scheme is only a few meters but can be reduced to less 1 m using 6-axis sensors [5]. The fingerprinting scheme estimates the relative position by matching real-time measurement parameters with the pre-measured location-based data [6]. This scheme is rather sensitive to the environment where features (i.e., fingerprints) of the measured data changes with the locations. The dead-reckoning is a fusing positioning method, which requires additional sensors (accelerometer or gyroscope) to estimate the relative position [7]. This scheme is susceptible to the accumulative PE therefore cannot provide long-term positioning accuracy. The triangulation method determines the absolute position by using the geometric properties of triangles [8]. In this scheme, the time of arrival (TOA), time difference of arrival (TDOA), and received signal strength (RSS) are used to estimate the position based on measuring distances from the transmitters (Txs), i.e., LEDs, to the Rx (mobile device). Whereas the angle of arrival (AOA) is adopted to measure angles relative to multiple LEDs to determine the position of the Rx [9-10]. However, TOA and TDOA are not feasible for real-time implementation because of difficulty in measuring the time difference as well as the need for perfect synchronization between the Tx and the Rx [9]. In vision analysis [11], adopted in augmented reality, the geometric relations between the objects' 3D positions in the real world and their 2D positions on their projections on the image sensor are used to determine the position. This is also known as P*n*P problem, which can be classified as the iterative or non-iterative method that are efficient, but suffer from instability for $n < 5$ [12]. A number of methods have been proposed for solving P*n*P for IPS, however, the existing of non-iterative solutions for $n < 5$ tend to be unstable in real implementation scenarios because of lack of redundant information [13]. In [14] this problem was resolved by introducing additional redundant points. However, this is not suitable for IPS using optical camera communication, since the distance between LEDs is fixed in deployment scenario and is difficult for a camera with a fixed-focal length to capture more LEDs in each frame.

In this letter, we present a real-time VLC-IPS based on a novel robust non-iterative P*n*P solution. In this scheme, the OOK data format is used for transmission of ID information of LEDs, which is detected and decoded using a smartphone camera with rolling shutter mechanism. The proposed scheme can determine the real-time camera's position and offer improved performance for $n = 4$ case, and is experimentally implemented and evaluated using a dedicated test bed (i.e., 100×100×80 cm^3). The rest of the letter is organized as

This work was supported in part by the Chunmiao Project of Haixi Institutes, Chinese Academy of Sciences, in part by the National Science Foundation of China under Grant 61601439 and 61501427, in part by State Key Laboratory of Advanced Optical Communication Systems and Networks, China.

Yiwei Li is with Institute of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan. (e-mail: lywei0306@foxmail.com)

Zabih Ghassemlooy is with Optical Communications Research Group, NCRLab, Faculty of Engineering and Environment, Northumbria University, Newcastle, UK, and College of Engineering, Huaqiao University, Quanzhou, China. (e-mail: z.ghassemlooy@northumbria.ac.uk)

Xuan Tang, Bangjiang Lin, Yi Zhang are with Quanzhou Institute of Equipment Manufacturing, Haixi Institutes, Chinese Academy of Sciences, Quanzhou, China. (e-mail: xtang@fjirsm.ac.cn; linbangjiang@163.com; yzhang.cn@outlook.com)

follows. Section II describes the proposed positioning algorithm in detail. Section III presents the experimental setup and performance evaluation for IPS, followed by the concluding remarks in Section IV.

II. Proposed Positioning Algorithm

Fig. 1 shows the schematic block diagram of the proposed scheme with the projections of 4 reference LEDs, which are uniformly distributed on the ceiling. Note that, only two of LEDs were chosen to transmit their unique coordinate information (UCI). Assuming that the world coordinates of LEDs and their corresponding 2D projections on the image plane are P_i^w and p_i^c (i=1, 2, 3, 4), respectively. The coordinate of the optical center O_c is considered as the position of the camera. Since using 2- and 3- LED will lead to multiple solutions, a 4-LED array was adopted to obtain a robust and unique solution with reduced complexity. Based on the PnP model given in [15], the coordinate of the LED in the image coordinate plane is given by:

$$X_c = \mathbf{R}X_w + \boldsymbol{t}\,, \tag{1}$$

where **R** and ***t*** denote the rotation matrix and translation vector, respectively, and X_w is the world coordinate of the LEDs in the world coordinate plane.

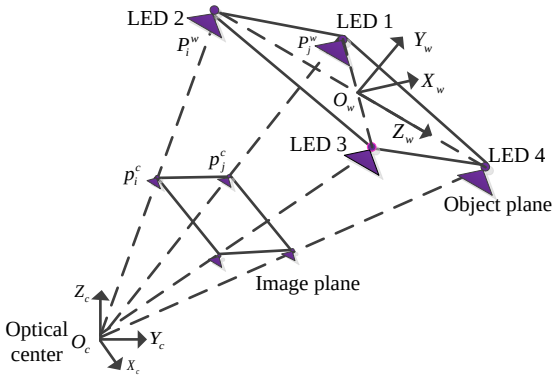

Fig. 1 A system block diagram of proposed IPS with LEDs.

As shown in Fig. 1, the LED array is projected onto the image plane as p_i^c. The edges of the reference LEDs are given by $\left\{\overrightarrow{P_i^w P_j^w} \mid i < j, i, j \in \{1,2,3,4\}\right\}$. First, an arbitrary edge $\overrightarrow{P_i^w P_j^w}$ is selected from this set as the rotation axis, based on which a new world coordinate system of $O_w X_w Y_w Z_w$ is created. The origin of $O_w X_w Y_w Z_w$ is at the center of $\overrightarrow{P_i^w P_j^w}$, and the direction of its Z_w-axis is the same as the vector itself, i.e., $\overrightarrow{P_i^w P_j^w}$. After the initial world coordinate system is converted into a $O_w X_w Y_w Z_w$, the transformation from $O_w X_w Y_w Z_w$ to camera coordinate system $O_c X_c Y_c Z_c$ is determined by using the rotation angle α around the Z_w-axis and the translation vector of $\overrightarrow{O_c O_w}$. Therefore, the crucial step in our method is to determinate the appropriate Z_w-axis.

In our system, we have used only two LEDs to transmit their UCIs. Note that, the world coordinate for all LEDs can be determined by knowing the coordinates of two LEDs based on the LED array's geometrical relationship. The two unmodulated LEDs at the diagonal position with higher light intensities are selected as the rotation axis of $O_w X_w Y_w Z_w$, see Fig. 1. Note that, LEDs with higher intensities offer higher signal to noise ratio in the presence of the ambient light and the longer edge is selected as Z_w-axis can be less affected by background illumination noise [14].

Following determination of the rotation matrix **R** from $O_w X_w Y_w Z_w$ to $O_c X_c Y_c Z_c$ is given by:

$$\begin{aligned}\mathbf{R} &= rot(X_w,\alpha)\cdot rot(Y_w,\beta)\cdot rot(Z_w,\gamma)\\ &= \begin{bmatrix}1 & 0 & 0\\ 0 & \cos\alpha & \sin\alpha\\ 0 & -\sin\alpha & \cos\alpha\end{bmatrix}\begin{bmatrix}\cos\beta & 0 & -\sin\beta\\ 0 & 1 & 0\\ \sin\beta & 0 & \cos\beta\end{bmatrix}\begin{bmatrix}\cos\gamma & -\sin\gamma & 0\\ \sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{bmatrix} \quad (2)\\ &= \begin{bmatrix}r_{11} & r_{12} & r_{13}\\ r_{21} & r_{22} & r_{23}\\ r_{31} & r_{32} & r_{33}\end{bmatrix}\begin{bmatrix}\cos\gamma & -\sin\gamma & 0\\ \sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{bmatrix},\end{aligned}$$

where α, β and γ are the deviation angles, which are measured from the world coordinate system to the image plane that correspond to the X_c, Y_c, Z_c- axes, respectively; $rot(X_w,\alpha)$ and $rot(Y_w,\beta)$ denote rotation of α and β around the X- and Y- axes, respectively. The rotation matrix $rot(X,\alpha)\cdot rot(Y,\beta)$ with 3rd column $\begin{bmatrix}r_{13} & r_{23} & r_{33}\end{bmatrix}^T = Z_w$ must meet the orthogonal constraint. By rotating the camera from $O_w X_w Y_w Z_w$ to $O_c X_c Y_c Z_c$, the center of Z_w is projected onto the center of the image plane. **R** can then be estimated by extracting the rotation angle γ. The other unknown parameters in (2) can be determined by means of orthogonal matrix constraint. The projection for one LED from the object plane to the image plane can be expressed as:

$$\begin{bmatrix}x_i^c\\ y_i^c\\ 1\end{bmatrix} = \rho_i\begin{bmatrix}r_{11} & r_{12} & r_{13}\\ r_{21} & r_{22} & r_{23}\\ r_{31} & r_{32} & r_{33}\end{bmatrix}\begin{bmatrix}\cos\gamma & -\sin\gamma & 0\\ \sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}X_i^w\\ Y_i^w\\ Z_i^w\end{bmatrix} + \begin{bmatrix}t_x\\ t_y\\ t_z\end{bmatrix}, \tag{3}$$

where ρ_i is a scaling factor for LED's UCIs and its projection to the image plane. $\begin{bmatrix}x_i^c & y_i^c & 1\end{bmatrix}^T$ is the coordinate of image point and $\begin{bmatrix}t_x & t_y & t_z\end{bmatrix}^T$ is the unknown translation vector.

Since there are five unknown parameters of $\cos\gamma$, $\sin\gamma$ and $\begin{bmatrix}t_x & t_y & t_z\end{bmatrix}^T$, at least five equations are required. The last row of (3) implies a linear relationship between γ and t_z. Substituting this expression in the first two rows, yields two linear independent equations for each LED. Therefore, with 4 LEDs, a total of 8 linear independent equations are considered as part of the proposed algorithm. With reference to (3), the UCI per LED and their corresponding projection on the image plane can be arranged into a 3×6 homogenous linear system. Therefore, by concatenating all 4 LEDs, a 12×6 homogenous linear system is developed in the form of $\mathbf{M}\boldsymbol{x} = 0$, where **M** and ***x*** can be expressed as:

$$\mathbf{M}=\begin{bmatrix} x_i r_{11}+y_i r_{12} & x_i r_{12}-y_i r_{11} & 1 & 0 & 0 & z_i r_{13}-x_i^c \\ x_i r_{21}+y_i r_{22} & x_i r_{22}-y_i r_{21} & 0 & 1 & 0 & z_i r_{23}-y_i^c \\ x_i r_{31}+y_i r_{32} & x_i r_{32}-y_i r_{31} & 0 & 0 & 1 & z_i r_{33}-1 \end{bmatrix} \quad (4)$$

and

$$\boldsymbol{x}=\begin{bmatrix}\cos\gamma & \sin\gamma & t_x & t_y & t_z & 1\end{bmatrix}^T. \quad (5)$$

The unknown variables vector $\boldsymbol{x}$ can be determined using the singular value decomposition (SVD) method as given by [12]. Our algorithm is summarized in Algorithm 1.

Algorithm 1 The proposed IPS Algorithm

1: Given world coordinate and image coordinate of 4 LEDs.
2: **for** $i=\{1,2,3,4\}$, $j=\{1,2,3,4\}$ do
3: Compute the $\left\|P_i^w P_j^w\right\|_2$ and the intensity of LED *i* and *j*;
4: **end for**
5: Obtain the maximum value of $\left\|P_i^w P_j^w\right\|_2$ and determine the rotation axis Z_w;
6: Compute the rotation matrix **R** and translation vector $\boldsymbol{t}$ using SVD method;
7: Output $-inv(\mathbf{R})\cdot\boldsymbol{t}$ as the position of camera.

III. Experiment Setup and Performance Evaluation

In this section, we demonstrate the performance of the proposed algorithm by means of experimental investigation as shown in Fig. 2, where the $O_w^{'}X_w^{'}Y_w^{'}Z_w^{'}$ is the initial world coordinate system. The initial world coordinates of LEDs are assigned as (40, 40, *h*), (40, -40, *h*), (-40, 40, *h*), and (-40, -40, *h*) for LEDs 1, 2, 3 and 4, respectively, where *h* is the height with values in the range of 20 cm to 80 cm. The LEDs optical foot-prints (i.e., coverage common area) project on the floor area covered by 4 LEDs representing a typical femtocell with a few meters of diameter.

At the Tx side, the STM8 chip was used to transmit IDs of LEDs in the OOK format via intensity modulation of the light sources. Here we used frequencies of 250 Hz and 500 Hz for IDs of LED 1 and LED 3, respectively. The OOK signal with different frequencies is generated by the STM8 TIM module, which is a 16-bit advanced control timer operating on counter mode. At the Rx side, the calibrated Meizu3 smartphone rear camera with an inherent resolution of 1920 × 1080 pixels was used to record a real-time video stream. The compressed real-time video stream was wirelessly transmitted to the computer using the installed ipcamera app on the smartphone. We set the resolution of video steam to 480 × 320 pixels on app to reduce the transmission delay. Next, the transmitted IDs were decoded by extracting the dark and bright stripes by means of image processing. Here, the dark and bright stripes of modulated LEDs were obtained using the Matlab image acquisition toolbox and the barcode detection algorithm [16]. The preset database was used to specify the relationship between the IDs and the initial world coordinates of the modulated LEDs (i.e., identify the initial world coordinates).

The decoding process of ID information can be separated into the following steps: (*i*) completing the image gray processing and binarization processing; (*ii*) image filtering using a binary OTSU filter [17]; (*iii*) detecting the dark and bright stripes using the barcode detection algorithm; and (*iv*) determining the projection center of LED using hough-circle detection processing. The key parameters for the experimental system are listed in Table I.

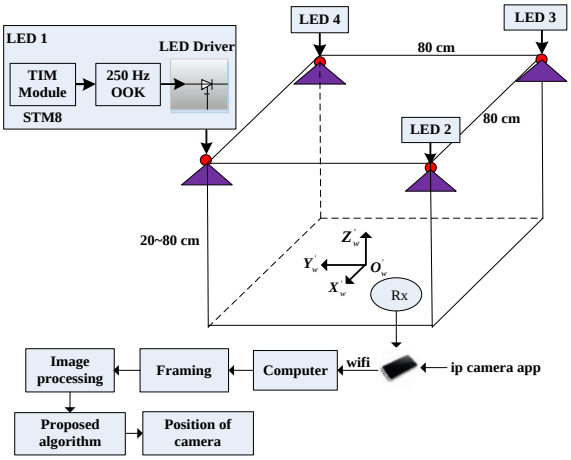

Fig. 2 Experimental setup for visible light positioning system (modules shown within LED1 also 500 Hz OOK applied to LED3).

Table I. System parameters

Parameter	Value
LED	NVC NLED9304
Transmit power	9 W
Camera	Meizu3 rear camera
Video resolution	480 ×320 pixels
Shutter speed	4 kHz
Camera aperture	f/24
Focal length	35 mm
Frame rate	10~30 fps

In order to evaluate the performance of the proposed VLC-IPS, we recorded a 10 seconds video and then determined the mean value for a single measurement. Fig. 3 illustrates the estimated and measured positions for a height of 50 cm with the mean and a maximum PE of 4.81 cm and 6.22 cm, respectively. From Fig. 3 we observe that PEs are much lower at the center of femtocell compared to the femtocell edges. This may be due to the object recognition of LEDs, where lens distortion results in inaccurate image coordinates of LEDs when the smartphone moves to femtocell edges. The distribution of the mean of PE for the height of 80 cm is depicted in Fig. 4. Comparing Fig. 3 and Fig. 4, we notice that the mean PE is increasing with the height.

Furthermore, when the smartphone moves outside the femtocell (i.e., capturing less than 4 LEDs in each frame), the system would not function as expected since at least 4 LEDs are required in the proposed algorithm. The proposed scheme can be used in larger rooms provided more frequencies representing increased number of LED IDs. Note that, when the smartphone moves between the femtocells, the PEs may not be the same in different femtocells because of the different ambient light levels and the field-of-view of the lens. Table II shows a comparison of the PEs for different schemes as well as the proposed system. The TOA and TDOA display lower PEs but cannot offer both real-time positioning and the height

estimation. The proposed system offers PEs similar to the image method [18] but with real-time 3D positioning.

Table II. Comparison of the VLC based Positioning system

Works	Methods used	PE	Room size	Positioning type
Ref. [4]	Proximity (Online)	The resolution of the LED grid	A large room	2D
Ref. [6]	Fingerprints (Online)	1 m	15×15×3 m	2D
Ref. [8]	RSS (Offline)	several cm	2×2×2.5 m	3D
Ref. [9]	TOA (Offline)	several cm	Height = 2.5 m	2D
Ref. [10]	TDOA (Offline)	1 cm	5×5×3 m	2D
Ref. [18]	Image (Offline)	~6 cm	0.5×0.5×1.2 m	2D
Proposed Scheme	Computer vision (Online)	~6 cm	1×1×0.8 m	3D

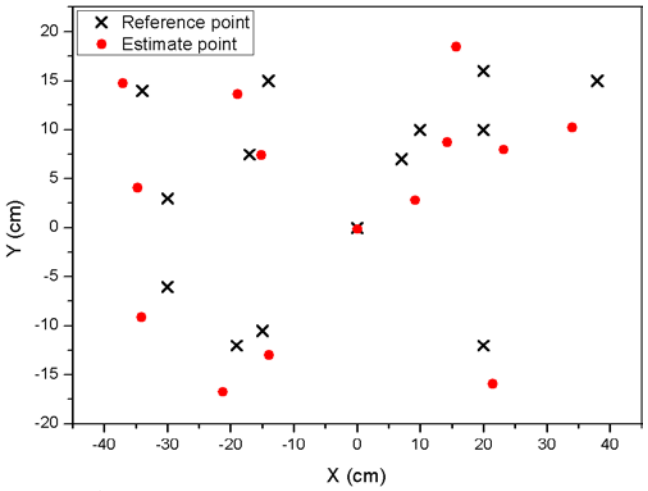

Fig. 3 The final estimated positions based on the proposed algorithm with height of 50 cm.

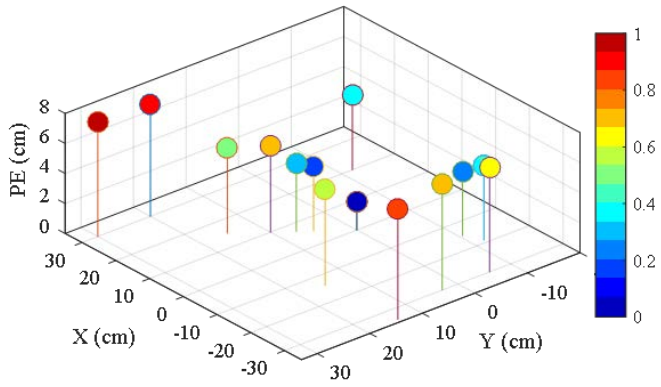

Fig. 4 The distribution of mean PEs with height of 80 cm.

IV. Conclusion

In this letter, we presented a VLC based real-time IPS using smartphone camera. The smartphone rolling shutter camera was used to decode the ID from the pixels occupied by the transmitting LEDs. The camera's position was robustly and uniquely determined using the proposed scheme. Experimental results showed that the proposed indoor positioning algorithm offered the mean PEs of 4.81 cm and 6.58 cm for the heights of 50 cm and 80 cm, respectively.

References

[1] Z. Ghassemlooy, S. Arnon, M. Uysal, Z. Xu and J. Cheng, "Emerging optical wireless communications-advances and challenges," *IEEE J. Sel. Areas Commun*, vol. 33, no. 9, pp. 1738-1749, Sept. 2015.

[2] Do, Trong-Hop, and Myungsik Yoo, "Potentialities and challenges of VLC based outdoor positioning," *Int. Conf. Info. Network* , Cambodia, 2015, pp. 474-477.

[3] B. Xie, K. Chen, G. Tan, et al. "LIPS: A light intensity--based positioning system for indoor environments". *ACM Trans. Sensor Networks*, vol. 12, no. 4, pp. 28-54, Sept. 2016.

[4] Y. U Lee, Kavehrad M, "Two hybrid positioning system design techniques with lighting LEDs and ad-hoc wireless network," *IEEE Trans. Consumer Electron*, vol.58, no. 4, pp. 1176-1184, Nov. 2012.

[5] Sertthin C, Ohtsuki T, Nakagawa M. 6-axis sensor assisted low complexity high accuracy-visible light communication based indoor positioning system," *IEICE Trans. Commun.*, vol. 93, no. 11, pp. 2879-2891, Nov. 2010.

[6] S. Feng, X. Li, R. Zhang, M. Jiang and L. Hanzo, "Hybrid positioning aided amorphous-cell assisted user-centric visible light downlink techniques," *IEEE Access*, vol. 4, pp. 2705-2713, May. 2016.

[7] Z. Li, L. Feng, A. Yang, "Fusion based on visible light positioning and inertial navigation using extended Kalman filters," *Sensors*, vol. 17, no. 5, pp. 1093-1103, Mar. 2017.

[8] S. H. Yang, H. S. Kim, Y. H. Son and S. K. Han, "Three-dimensional visible light indoor localization using AOA and RSS with multiple optical receivers," *J. Lightw. Technol.*, vol. 32, no. 14, pp. 2480-2485, July. 2014.

[9] T. Q. Wang, Y. A. Sekercioglu, A. Neild and J. Armstrong, "Position accuracy of time-of-arrival based ranging using visible light with application in indoor localization systems," *J. Lightw. Technol.*, vol. 31, no. 20, pp. 3302-3308, Oct. 2013.

[10] S. Y. Jung, S. Hann and C. S. Park, "TDOA-based optical wireless indoor localization using LED ceiling lamps," *IEEE Trans. Consum. Electron.*, vol. 57, no. 4, pp. 1592-1597, Nov. 2011.

[11] Y. Gu, Lo. A, Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Commun Surv Tut.*, vol. 11, no. 1, pp. 13-32, Mar. 2009.

[12] S. Li, C. Xu and M. Xie, "A robust $O(n)$ solution to the perspective-n-point problem," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1444-1450, July, 2012.

[13] A. Penate-Sanchez, J. Andrade-Cetto and F. Moreno-Noguer, "Exhaustive linearization for robust camera pose and focal length estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 10, pp. 2387-2400, Oct. 2013.

[14] S. Li, C Xu. "Efficient lookup table based camera pose estimation for augmented reality." *Computer animation and virtual worlds*, vol. 22, no. 1, pp. 47-58, Jan. 2011.

[15] Q. Zhang, Y Zhang, Z. Tian, "The method of solving the non-coplanar perspective-four-point (P4P) problem," *in Proc. 33rd Chinese Control Conf.*, 2014, pp. 1039-1043.

[16] C. Fu, C. W. Cheng, W. H. Shen, Y. L. Wei and H. M. Tsai, "Lightbib: marathoner recognition system with visible light communications," *IEEE Int. Conf. Data Science and Data Intensive Sys.*, 2015, pp. 572-578.

[17] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62-66, 1979.

[18] 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.