Novel Visible Light Communication Assisted Perspective-Four-Line Algorithm for Indoor Localization

Lin Bai†, Yang Yang†, Chunyan Feng†, Caili Guo‡ and Bowen Jia†

Chinese University of Technology and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
Emails: {bailin2126, young0607, cyfeng}@bupt.edu.cn
†Beijing Laboratory of Advanced Information Networks, School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
Email: guocaili@bupt.edu.cn
‡China Electronics Standardization Institute, Beijing 100176, China

Abstract — In this paper, a novel visible light communication (VLC) assisted Perspective-four-Line algorithm (V-P4L) is proposed for indoor localization. The basic idea of V-P4L is to joint the coordinate information obtained by VLC and the geometric information in computer vision for practical indoor localization. In particular, V-P4L first exploits the geometric information to estimate the orientation and coordinate of a single rectangular LED luminaire in the camera coordinate system based on the plane and solid geometry. Then, VLC is used to transmit the world coordinate information of the luminaire to the camera. Next, the position and pose of the camera in the world coordinate system are obtained by the single-view geometry and the linear least square method. Due to the combination of VLC and computer vision, V-P4L only requires a single luminaire for localization and does not rely on the channel model. Therefore, V-P4L is more robust to dust and other factors that interfere visible light signals. Besides, unlike conventional Perspective-n-Line (PnP) algorithms, V-P4L does not require the 3 dimensional (3D)-2 dimensional (2D) correspondences. Simulation results show that for V-P4L the location error is always less than $18\,\text{cm}$ and the orientation error is always less than $4.5^\circ$.

I. INTRODUCTION

Accurate indoor localization is increasingly important due to the surging location based services such as location tracking, navigation and robot movement control. In this research field, visible light positioning (VLP) and computer vision based localization have gained increasing attentions due to their high accuracy and low cost [1], [2].

Typical VLP algorithms include proximity [3], fingerprinting [4], [5], time of arrival (TOA) [6], angle of arrival (AOA) [7], received signal strength (RSS) [8] and image sensing [9]. Among these algorithms, only image sensing can estimate the pose of the receiver. Besides, image sensing, TOA, AOA and RSS require multiple luminaires for localization [1], which is infeasible when the available luminaires are insufficient. On the other hand, proximity and fingerprinting, which cannot estimate the receiver pose, can estimate the receiver position using a single luminaire. However, the accuracy of proximity

is insufficient. Besides, in fingerprinting, ambiguity issues are serious, and thus accuracy is low [10]. It has been justified that at least three luminaires are required in fingerprinting to reduce the effect of ambiguity [10]. In addition, RSS algorithms are most widely-used due to their high accuracy and low cost [8]. However, RSS algorithms rely on accurate channel model, which is the most serious problem for them. A popular assumption in RSS algorithms is that the radiation pattern of LEDs is the Lambertian model which may not be true for many luminaires especially when a lampshade is used [11]. Meanwhile, the estimated channel gain is affected by sunlight, dust and shadowing in practice [12], [13].

On the other hand, typical computer vision based localization methods include Perspective-n-Line (PnP) and Perspective-n-Point (PnP). PnP methods perform by analyzing $n$ correspondences between 3 dimensional (3D) reference lines and their 2 dimensional (2D) projections (i.e., 3D-2D correspondences). Compared with PnP methods which exploit the point features, PnP methods achieve higher detection accuracy and are more robust to (partial) occlusions [14], [15]. However, PnP methods need 3D-2D correspondences which are difficult to obtain. In existing PnP studies, the 3D-2D correspondences are given directly, which is impractical in practice [2]. To circumvent this challenge, the work in [2] proposed a method to find the 3D-2D correspondences for the scenario where the vertical lines are more than horizontal lines. However, this method cannot be applied to the scenario where the rectangular beacons are deployed on the ceiling since there is no significant difference between numbers of lines of different directions.

The main contribution of this paper is to propose a novel visible light communication (VLC) assisted Perspective-four-Line algorithm (V-P4L). To the authors’ best knowledge, this is the first VLP algorithm that only requires a single luminaire for position and pose estimation. The basic principle of the scheme is to joint the coordinate information obtained by VLC and the geometric information in computer vision to

†In this paper, a luminaire means that all the LEDs in a luminaire transmit the same information.
calculate the absolute position and pose of the camera based on the single-view geometry and the linear least square (LLS) method. Therefore, compared with RSS algorithms, V-P4L only requires a single luminaire and does not rely on the ideal assumption of channel model. Besides, compared with PnL methods, V-P4L does not require the 3D-2D correspondences due to use of VLC. Note that V-P4L requires a VLC-enabled rectangular LED luminaire as the beacon. Fortunately, rectangular LED luminaires including rectangular LED array and the rectangular luminaires constructed by LED tubes are popular in indoor scenarios. Simulation results show that for V-P4L the location error is always less than 18 cm and the orientation error is always less than 4.5°.

The rest of the paper is organized as follows. Section II introduces the system model. The proposed localization algorithm is detailed in Section III. Simulation results are presented in Section IV. Finally, the paper is concluded in Section V.

II. SYSTEM MODEL

The system diagram is illustrated in Fig. 1. Four coordinate systems are utilized for pixelization, which are the pixel coordinate system (PCS) \( o^P - u^P v^P \) on the image plane, the image coordinate system (ICS) \( o^I - x^I y^I \) on the image plane, the camera coordinate system (CCS) \( o^C - x^C y^C z^C \) and the world coordinate system (WCS) \( o^W - x^W y^W z^W \). In PCS, ICS, and CCS, the axes \( u^P, x^I \) and \( x^C \) are parallel to each other and, similarly, \( v^P, y^I \) and \( y^C \) are also parallel to each other. Besides, \( o^P \) and \( o^I \) are on the same straight line. The distance between \( o^P \) and \( o^I \) is the focal length \( f \), and thus the \( z \)-coordinate of the image plane in CCS is \( z^C = f \). In addition, \( o^C \) is in the upper left corner of the image plane.

A VLC-enabled rectangular LED luminaire constructed by four vertices \( P_i \ (i \in \{1, 2, 3, 4\}) \) is the transmitter mounted on the ceiling. The four 3D reference lines \( L_{ij} \ (i, j \in \{1, 2, 3, 4\}, i \neq j \) are the edges of the luminaire. Without loss of generality, the luminaire is assumed to face vertically downwards. Therefore, the unit normal vector of the luminaire in WCS, \( \mathbf{n}_{LED} = (0, 0, 1)^T \), where \( (\cdot)^T \) denotes the transpose of matrices, is known in advance. Besides, \( s_i^w = (x_i^w, y_i^w, z_i^w)^T \) is the coordinate of the \( i \)-th vertex of the luminaire in WCS, which are assumed to be known at the transmitter and can be obtained by the receiver through VLC [16], [17]. Note that all the LEDs in the luminaire transmit the same information.

On the other hand, the receiver is a standard pinhole camera which is not coplanar with the luminaire. Therefore, the transmitter and the receiver produced a rectangular pyramid \( \delta^P - P_1 P_2 P_3 P_4 \). In the rectangular pyramid \( \delta^P - P_1 P_2 P_3 P_4 \), the rectangle \( P_1 P_2 P_3 P_4 \) is called the base face. Meanwhile, we define \( \Pi_{ij} \) as the lateral face determined by the vertices \( P_i, P_j \ (i, j \in \{1, 2, 3, 4\}, i \neq j \) and \( \delta^P \). Besides, \( P_i \) is the \( i \)-th vertex, \( \delta^P \) is called the apex. In the camera, \( p_i \) is the projection of the \( i \)-th vertex of the LED luminaire \( P_i \) on the image plane. Besides, \( l_{ij} \) is the 2D projection on the image plane of \( L_{ij} \). Note that many existing PnL algorithms assume that the 3D-2D correspondences \( (L_{ij} \Leftrightarrow l_{ij}) \) are known in advance, which is too ideal in practice [2]. In contrast, in this work, the 3D-2D correspondences are unknown.

The pixel coordinate of \( p_i \) is denoted by \( s_i^p = (u_i^p, v_i^p)^T \), and this coordinate can be obtained by the camera through image processing [9]. Based on the single-view geometry, the image coordinate of \( p_i \), \( s_i^I = (x_i^I, y_i^I)^T \), can be obtained as follows

\[
\begin{align*}
(x_i^I &= (u_i^p - u_0) d_x \\
y_i^I &= (v_i^p - v_0) d_y
\end{align*}
\]

where \( d_x \) and \( d_y \) are the physical size of each pixel in the \( x \) and \( y \) directions on the image plane, respectively. Besides, \( (u_0, v_0)^T \) is the pixel coordinate of \( o^I \) which is termed as the principal point. The camera’s intrinsic parameters, including \( (u_0, v_0)^T \) and the focal ratio \( f_u = \frac{f}{2} \) and \( f_v = \frac{f}{2} \), can be calibrated in advance [8]. The transformation from CCS to WCS can be expressed as follows

\[
s^w = R^w_c s^C + T^w_c
\]

where \( s^w \) and \( s^C \) are the world and camera coordinates of the same object, respectively. Besides, \( R^w_c \) and \( T^w_c \) denote the pose and the position of the camera in WCS, respectively. The task of the localization is to find out \( R^w_c \) and \( T^w_c \).

III. VISIBLE LIGHT COMMUNICATION ASSISTED PERSPECTIVE-FOUR-LINE ALGORITHM (V-P4L)

In this section, a novel localization algorithm, termed as V-P4L is proposed. V-P4L mainly consists of 3 steps. In the first step, the normal vector of the LED luminaire in CCS, \( \mathbf{n}_{LED} \), is estimated based on the single-view geometry and the plane and solid geometry. Then, based on \( \mathbf{n}_{LED} \), the camera coordinates of the four luminaire’s vertices \( s_i^w, i \in \{1, 2, 3, 4\} \), are estimated by the solid geometry. Finally, based on \( \mathbf{n}_{LED} \), \( s_i^w \) and \( s_i^w \), the pose and position of the camera in WCS can be estimated by the single-view geometry and the LLS algorithm.

A. The Normal Vector Of The Luminaire In CCS

In ICS, the point-normal form equation of a given \( l_{ij} \) can be expressed as \( x^I \cos \phi_{ij} + y^I \sin \phi_{ij} = \rho_{ij} \) [18], where \( (x^I, y^I)^T \)
is the image coordinate of a point on \( l_{ij} \), \( \phi_{ij} \) is the rotate angle from \( y' \)-axis to \( l_{ij} \) in anticlockwise direction and \( \rho_{ij} \) is the distance from \( o' \) to \( l_{ij} \). Since \( p_i \) and \( p_j \) are on \( l_{ij} \), \( \phi_{ij} \) and \( \rho_{ij} \) can be obtained easily. Let \( \mathbf{n}^c = (A, B, C)^T \). In CCS, let \( \mathbf{n}_{ij}^c = (A_{ij}, B_{ij}, C_{ij})^T \) \( (i, j \in \{1, 2, 3, 4\}, i \neq j) \) denote the normal vector of \( \Pi_{ij} \) and \( \mathbf{N}_{ij}^c \in \mathbb{R}^3 \) \( (i, j \in \{1, 2, 3, 4\}, i \neq j) \) denote the direction vector of \( l_{ij} \). Since \( L_{ij} \) is the intersection line of the plane of \( P_1 P_2 P_3 P_4 \) and \( \Pi_{ij} \), \( \mathbf{N}_{ij}^c \) can be calculated as following:

\[
\mathbf{N}_{ij}^c = \mathbf{n}^c \times \mathbf{n}_{ij}^c.
\]

Based on the solid geometry, we have

\[
\begin{align*}
N_{131} \cdot n_{12} &= 0 \quad \text{and} \\
N_{111} \cdot n_{21} &= 0.
\end{align*}
\]

Define \( m = \frac{a}{c} \) and \( n = \frac{b}{c} \), and we can obtain \( m \) and \( n \) as the functions of \( A_{ij}, B_{ij} \) and \( C_{ij} \) by solving (5). Therefore, the normalized normal vector of the rectangle \( P_1 P_2 P_3 P_4 \) (i.e., the orientation of the luminaire) in CCS can be expressed as

\[
\mathbf{n}_{\text{LED}} = (\cos \alpha, \cos \beta, \cos \gamma)^T
\]

where

\[
\begin{align*}
\cos \alpha &= \frac{m}{\sqrt{m^2 + n^2 + 1}} \\
\cos \beta &= \frac{n}{\sqrt{m^2 + n^2 + 1}} \\
\cos \gamma &= \frac{1}{\sqrt{m^2 + n^2 + 1}}
\end{align*}
\]

B. Camera Coordinates Of The Luminaire’s Vertices

Since \( P_3 \) is the intersection point of the rectangle \( P_1 P_2 P_3 P_4 \), \( \Pi_{12} \) and \( \Pi_{11} \), the camera coordinate of \( s_i^c = (x_i^c, y_i^c, z_i^c)^T \) can be calculated as following:

\[
s_i^c = W_i s_i.
\]

In general, the camera coordinate

\[
\begin{bmatrix}
A_{12} & B_{12} & C_{12} \\
A_{41} & B_{41} & C_{41}
\end{bmatrix}
\]

\[
\begin{bmatrix}
m & n & 1 \\
1 & 0 & 0
\end{bmatrix}
\]

The volume of the rectangular pyramid \( V = \frac{1}{6} P_1 P_2 P_3 P_4 \) can be calculated as follows

\[
V = \frac{1}{2} \sum_{i=1}^{4} V_i.
\]

For the triangular pyramid \( \Pi_{12} \), its volume can be calculated as follows

\[
V_1 = \frac{1}{6} \det(M_{11})
\]

where \( \det(M) \) is the determinant of the matrix \( M \) and

\[
M_{11} = \begin{bmatrix}
x_1^c & y_1^c & z_1^c \\
x_2^c & y_2^c & z_2^c \\
x_3^c & y_3^c & z_3^c
\end{bmatrix}.
\]

Substituting (8) into (9), we can have

\[
V_1 = \frac{g_1}{6},
\]

\[
\text{where } g_1 = \frac{1}{6} \det(M_{12}),
\]

\[
M_{12} = \begin{bmatrix}
W_1 & W_2 & W_3
\end{bmatrix}^T.
\]

The volumes of the other three triangular pyramid \( \Pi_{12}, \Pi_{13}, \Pi_{14} \), \( \Pi_{13} \), and \( \Pi_{14} \), respectively, can be obtained in the same way. Since \( V = \frac{1}{2} \sum_{i=1}^{4} V_i \), \( C \) can be calculated as follows

\[
C = \sqrt{3} \sum_{i=1}^{4} \sqrt{m^2 + n^2 + 1}.
\]

Substituting (10) into (8), the camera coordinate of \( P_i \) \( (i \in \{1, 2, 3, 4\}) \) can be obtained.

C. Pose And Position Of The Camera In WCS

Based on the orientation and coordinate information of the luminaire, the position and pose of the camera in WCS can be obtained with the help of VLC. In particular, the position and pose estimation can be implemented in the following 3 steps.

1) Calculate the rotation angles corresponding to the \( x^c \)-axis and \( y^c \)-axis:

Let \( R_X, R_Y, R_Z \) denote the rotation matrices along the \( x^c \)-axis, \( y^c \)-axis and \( z^c \)-axis, respectively. Given \( R_X, R_Y, R_Z \) and \( R_Z \) as follows

\[
R_X = \begin{bmatrix}
1 & 0 & 0 \\
0 & \cos \varphi & \sin \varphi \\
0 & -\sin \varphi & \cos \varphi
\end{bmatrix},
\]

\[
R_Y = \begin{bmatrix}
\cos \theta & 0 & \sin \theta \\
0 & 1 & 0 \\
-\sin \theta & 0 & \cos \theta
\end{bmatrix},
\]

\[
R_Z = \begin{bmatrix}
\cos \psi & -\sin \psi & 0 \\
\sin \psi & \cos \psi & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

where \( \varphi, \theta, \psi \) are the unknown Euler angles corresponding to the \( x^c \)-axis, \( y^c \)-axis and \( z^c \)-axis, respectively.

The rotation angles \( \varphi, \theta \) corresponding to the \( x^c \)-axis and \( y^c \)-axis can be obtained by solving (15).

Based on the single-view geometry, the relationship between

\[
\mathbf{n}_{\text{LED}} = (0, 0, 1)^T
\]

and

\[
\mathbf{n}_{\text{LED}} = (\cos \alpha, \cos \beta, \cos \gamma)^T
\]

can be given as

\[
\mathbf{n}_{\text{LED}} = R^\alpha \cdot \mathbf{n}_{\text{LED}}^c.
\]

Therefore, we have

\[
\begin{align*}
\cos \alpha &= -\sin \theta \\
\cos \beta &= \cos \theta \cdot \sin \varphi \\
\cos \gamma &= \cos \theta \cdot \cos \varphi.
\end{align*}
\]
2) Calculate the z-coordinate of the camera in WCS:

Based on the camera coordinate of $P_i$ ($i \in \{1, 2, 3, 4\}$), the world coordinate of $P_i$ can be given as

$$ s_i^w = R_i^w s_i^c + T_i^w $$  \hspace{1cm} (16)$$

where $s_i^w = (x_i^w, y_i^w, z_i^w)^T$ is known in advance and can be obtained at the receiver via VLC. $s_i^c = (x_i^c, y_i^c, z_i^c)^T$ is calculated in subsection III-B and $T_i^c = (t_x, t_y, t_z)^T$ is the 3D world coordinate of the camera. Let $R_Y R_X = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$. We can rewrite (16) as follows

$$ \begin{bmatrix} x_i^w \\ y_i^w \\ z_i^w \end{bmatrix} = \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} x_i^c \\ y_i^c \\ z_i^c \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} $$  \hspace{1cm} (17)$$

In (17), there are four unknown parameters $\psi, t_x, t_y$ and $t_z$. If the 3D-2D correspondences are known in advance, we can easily obtain the four unknown parameters with the four vertices’ world and camera coordinates. However, in this work, the 3D-2D correspondences are unknown for practical considerations. Therefore we can only obtain the relationship between $z_i^w$ and $t_z$ from (17) as $z_i^w = (x_i^c \cdot c_1 + y_i^c \cdot c_2 + z_i^c \cdot c_3) + t_z$ due to the fact that all the LEDs are deployed on the ceiling at the same height, i.e., $z_i^w = z_i^w = z_i^w = z_i^w$. The z-coordinate of the camera in WCS $t_z$ can be calculated as follows

$$ t_z = \frac{1}{4} \left( \sum_{i=1}^{4} z_i^w - \sum_{i=1}^{4} (x_i^c \cdot c_1 + y_i^c \cdot c_2 + z_i^c \cdot c_3) \right) \hspace{1cm} (18)$$

3) Calculate $\psi, t_x$ and $t_y$:

There are still three unknown parameters $\psi, t_x$ and $t_y$ in (17). These unknown parameters can be calculated by the LLS estimator, which can be expressed in a matrix form as follows

$$ A_{ij} X = b_{ij} $$ \hspace{1cm} (19)$$

where $A_{ij} = [A_i, A_j]^T$ and $b_{ij} = [b_i, b_j]^T \ (i, j \in \{1, 2, 3, 4\}, i \neq j)$, where

$$ A_i = \begin{bmatrix} a_1 x_i^c + a_2 y_i^c + a_3 z_i^c & -b_1 x_i^c - b_2 y_i^c - b_3 z_i^c & 1 \\ b_1 x_i^c + b_2 y_i^c + b_3 z_i^c & a_1 x_i^c + a_2 y_i^c + a_3 z_i^c & 0 \\ 0 & 1 \end{bmatrix} $$ \hspace{1cm} (20)$$

$$ X = [\cos \psi, \sin \psi, t_x, t_y]^T $$ \hspace{1cm} (21)$$

and

$$ b_{ij} = [x_i^w, y_i^w]^T $$ \hspace{1cm} (22)$$

Therefore, the unknown parameters can be given by

$$ \hat{X} = (A_i^T A_i)^{-1} A_i^T b_{ij} $$ \hspace{1cm} (23)$$

where $\hat{X}$ is the estimate of $X$.

Since the 3D-2D correspondences are not known in advance, given a certain $A_{xy}$ and $b_{ij}$ ($r, s, i, j \in \{1, 2, 3, 4\}, r \neq s, i \neq j$), we do not know their exact correspondence relationship. Fortunately, there are four $A_x$ ($r \in \{1, 2, 3, 4\}$) and $b_i$ ($i \in \{1, 2, 3, 4\}$, and that means there are total $C_4^2 = 6$ different $A_{xy}$ and $b_{ij}$ ($r, s, i, j \in \{1, 2, 3, 4\}, r \neq s, i \neq j$).

Therefore, for each $b_{ij} = [b_i, b_j]^T \ (i, j \in \{1, 2, 3, 4\}, i \neq j)$, we can obtain 6 candidate solutions corresponding to 6 $A_{xy}$ ($r, s \in \{1, 2, 3, 4\}, r \neq s$), and one of which is $X_{ij}$, where $X_{ij}$ represents the exact $X$ corresponding to $b_{ij}$. When considering 6 different $b_{ij}$, we can obtain total 36 solutions which can further be separated into 6 groups. To improve the accuracy performance of V-P4L, the final estimated $X$ can be calculated by averaging the 6 closest solutions (i.e., 6 $X_{ij}$) in the 6 groups of solutions, which can be expressed as follows

$$ \hat{X} = \frac{1}{6} \sum_{i=1}^{4} \sum_{j=1, j \neq i}^{4} X_{ij} $$ \hspace{1cm} (24)$$

In this way, the position and pose of the camera, which are $R_i^w$ and $T_i^w$, respectively, can be obtained. Although V-P4L requires the LED luminaire to be a rectangle, V-P4L is robust to partial occlusion, which is meaningful for the moving cases. For instance, if the projection of $P_3$, $p_1$, is blocked by barriers and not on the image plane, the pixel coordinate of $p_1$ can be determined by the intersection of $l_{12}$ and $l_{41}$, and thus V-P4L can still successively implement.

IV. SIMULATION RESULTS AND ANALYSES

Since V-P4L combines VLC and computer vision based localization, a VLP algorithm named camera assisted received signal strength ratio algorithm (CA-RSSR) [8], and a typical computer vision algorithm termed P4L algorithm [18] are conducted as the baselines schemes in this section.

A. Basic Setup For Simulation

The system parameters are listed in Table I. The LED luminaire is deployed in the center of the ceiling. The length of the luminaire is set to 120 cm [4], [19], and the widths of luminaire are different in different simulation configurations. All statistical results are averaged over 10000 independent runs. For each simulation run, the receiver positions are selected in the room randomly. The pinhole camera is calibrated. Since the image noise affects the pixel coordinate of the luminaire’s projection on the image plane, the pixel coordinate is obtained by processing 20 images for the same position. Note that since CA-RSSR requires five LEDs for 3D localization, we assume that the LEDs at the vertices and the center of the luminaire are used for CA-RSSR. Besides, all the LEDs transmit different information in CA-RSSR, which is not required in V-P4L and the P4L algorithm. Therefore, compared with V-P4L,
the VLC link of CA-RSSR is more complex. Furthermore, CA-RSSR relies on the perfect Lambertian pattern model. However, the VLC channel model can be quiet different from the Lambertian pattern model even using the LED having nearly-ideal Lambertian pattern, as shown in [11], and the difference can be over 100% in certain cases. Therefore, we set a random deviation $\delta_1 \leq 10\%$ for the Lambertian pattern model for CA-RSSR, conservatively. On the other hand, the P4L algorithm exploits a rectangle to estimate the position and pose of the camera. The P4L algorithm assumes that the camera knows the 3D-2D correspondences. However, the beacon in the P4L algorithm cannot convey its coordinate information to the camera, which make the assumption impractical. Besides, the method to find the 3D-2D line correspondences given by [2] is also not practical when the camera captures the beacons on the ceiling as stated in Section I. Therefore, we set a random deviation $\delta_2 \leq 10\%$ for the 3D-2D correspondences, conservatively. In addition, the P4L algorithm can only obtain the relative position. For comparison with V-P4L, we transform the relative position into the absolute position for the P4L algorithm in this section.

We evaluate the performance of V-P4L in terms of its accuracy of position and pose estimation. We define location error as $LE = \|r_{true}^{w} - r_{est}^{w}\|$, where $r_{true}^{w} = (x_{true}^{w}, y_{true}^{w}, z_{true}^{w})$ and $r_{est}^{w} = (x_{est}^{w}, y_{est}^{w}, z_{est}^{w})$ are the actual and estimated world coordinates of the receiver, respectively. Besides, the orientation error is defined as $OE = |\Theta_{true} - \Theta_{est}|$, where $\Theta_{true}$ and $\Theta_{est}$ are the actual and estimated rotation angles, respectively.

### B. Performance Analysis

In this subsection, we evaluate the performance of V-P4L under various widths of the luminaire and the image noise.

We first evaluate the effect of the width of the luminaire on localization accuracy of V-P4L. This performance is represented by the means of location errors with the width varying from 20 cm to 100 cm. The image noise is modeled as a white Gaussian noise having an expectation of zero and a standard deviation of 2 pixels [20]. As shown in Fig. 2, V-P4L is able to obtain the best performance among the three algorithms. For V-P4L, the means of location errors keep below 23 cm for all the widths of the luminaire. In contrast, for CA-RSSR, the means of location errors decrease from 102 cm to 68 cm as the width of the luminaire increases from 20 cm to 100 cm. Besides, for the P4L algorithm, the means of location errors decrease from 105 cm to 38 cm. As shown in Fig. 2, the localization accuracy increases with the increase of the width for all the three algorithms. In addition, the location errors of V-P4L are always less than 18 cm using a single LED luminaire whose width is longer than 20 cm, and thus V-P4L can be applied to popular indoor luminaires.

Since the accuracy of pose estimation is also affected by the width of the LED luminaire, we then compare the orientation errors along the $x$-axis, $y$-axis and $z$-axis between V-P4L and the P4L algorithm with varying widths of the luminaire. As shown in Fig. 3, for both V-P4L and the P4L algorithm, the means of orientation errors along the $x$-axis and $y$-axis keep below $4^\circ$ for all the widths of the LED luminaire. Besides, for V-P4L, the means of orientation errors along the $z$-axis decrease from $4.5^\circ$ to $1.5^\circ$ as the width of the luminaire increases. In contrast, for the P4L algorithm, the means of orientation errors along the $z$-axis decrease from 10.2° to 6.8°. Therefore, compared with the P4L algorithm, V-P4L can obtain higher accuracy for pose estimation.

Then, we evaluate the effect of the image noise on the localization performance of V-P4L when the width of the luminaire is 40 cm. The image noise is modeled as a white Gaussian noise having an expectation of zero and a standard deviation, $\sigma_n$, ranging from 0 to 4 pixels [20]. Figure 4 shows the means of location errors versus image noises. As shown in Fig. 4, for V-P4L, the mean of location errors for $\sigma_n = 0$ pixel is 0 cm. That indicates the location errors of V-P4L is totally caused by the image noise. Besides, the means of location errors increase from 0 cm to 18 cm as the image noise increases from 0 to 4 pixels for V-P4L. In contrast, for the P4L algorithm, the means of location errors increase from 30 cm to 40 cm. In addition, for CA-RSSR, the means of location errors.
errors are around 90 cm. Therefore, compared with the P4L algorithm and CA-RSSR, V-P4L can obtain higher accuracy for position estimation.

Finally, we compare the accuracy of pose estimation between V-P4L and the P4L algorithm under different image noises. Figure 5 shows the means of orientation errors along x-axis, y-axis and z-axis with the image noise ranging from 0 to 4 pixels. As shown in Fig. 5, V-P4L performs consistently well from 0 to 4 pixels with the means of orientation errors along all x-axis, y-axis and z-axis below 3.5°. In contrast, for the P4L algorithm, the means of orientation errors along the z-axis increase from 6° to 8°. Therefore, compared with the P4L algorithm, V-P4L can obtain more stable and accurate pose estimation regardless of the image noise.

V. CONCLUSION

We have proposed a novel indoor localization algorithm named V-P4L that estimates the position and pose of the camera using a single, VLC enabled LED luminaire. V-P4L does not require ideal channel model assumption. Besides, based on the capability of VLC, V-P4L does not require the 3D-2D correspondences. Therefore, V-P4L can obtain higher accuracy performance for position and pose estimation than CA-RSSR and PnL algorithms. Simulation results have shown that for V-P4L the location error is always less than 18 cm and the orientation error is always less than 4.5°. In the future, we will experimentally implement V-P4L based on a dedicated test bed.

REFERENCES

[1] T. H. Do and M. Yoo, “An in-depth survey of visible light communication based positioning systems,” Sensors, vol. 16, no. 5, pp. 678, May 2016.
[2] T. Goto, S. Pathak, Y. Ji, H. Fujii, A. Yamashita, and H. Asama, “Line-based global localization of a spherical camera in manhattan worlds,” in 2018 IEEE Int. Conf. Robotics Automation (ICRA), 2018, pp. 2296–2303.
[3] C. Serthlin, T. Ohtsuki, and M. Nakagawa, “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.
[4] K. Qiu, F. Zhang, and L. Ming, “Let the light guide us: VLC-based localization,” IEEE Robot. Autom. Mag., vol. 23, no. 4, pp. 174–183, Dec. 2016.
[5] M. Chen, Z. Yang, W. Saad, C. Yin, H.V. Poor, and S. Cui, “A joint learning and communications framework for federated learning over wireless networks,” IEEE Trans. Wireless Commun., to appear, 2020.
[6] T. Q. Wang, Y. A. Sekercioğlu, 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.
[7] B. Zhu, J. Cheng, Y. Wang, J. Yan, and J. Wang, “Three-dimensional VLC positioning based on angle difference of arrival with arbitrary tilting angle of receiver,” IEEE J. Sel. Areas Commun., vol. 36, no. 1, pp. 8–22, Jan. 2018.
[8] L. Bai, Y. Yang, C. Guo, C. Feng, and X. Xu, “Camera assisted received signal strength ratio algorithm for indoor visible light positioning,” IEEE Commun. Lett., vol. 23, no. 11, pp. 2022–2025, Nov. 2019.
[9] 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.
[10] A. M. Vegni and M. Biagi, “An indoor localization algorithm in a small-cell LED-based lighting system,” in Proc. IEEE Int. Conf. Indoor Position. Indoor Navig. (IPIN), 2012, pp. 1–7.
[11] F. Miramirkhani and M. Uysal, “Channel modeling and characterization for visible light communications,” IEEE Photonics J., vol. 7, no. 6, pp. 1–16, Dec. 2015.
[12] Z. Dong, T. Shang, Y. Gao, and Q. Li, “Study on VLC channel modeling under random shadowing,” IEEE Photonics J., vol. 9, no. 6, pp. 1–16, Dec. 2017.
[13] A. Culean and M. Dimian, “Current challenges for visible light communications usage in vehicle applications: A survey,” IEEE Commun. Surveys Tuts., vol. 19, no. 4, pp. 2681–2703, 4th Quart. 2017.
[14] B. Přibyl, P. Zemčík, and M. Cadlák, “Absolute pose estimation from line correspondences using direct linear transformation,” Comput. Vis. Image Und., vol. 161, pp. 130–144, Aug. 2017.
[15] Alexander V., Jan F., and Franscesc M.-N., “Accurate and linear time pose estimation from points and lines,” in Europ. Conf. Comp. Visi. (ECCV), 2016, pp. 583–599.
[16] Y. Yang, Z. Zeng, J. Cheng, and C. Guo, “An enhanced DCO-OFDM scheme for dimming control in visible light communication systems,” IEEE Photon. J., vol. 8, no. 3, pp. 1–13, Jun. 2016.
[17] Y. Yang, Z. Zeng, J. Cheng, C. Guo, and C. Feng, “A relay-assisted OFDM system for VLC uplink transmission,” IEEE Trans. Commun., vol. 67, no. 9, pp. 6268–6281, Sept. 2019.
[18] Wang X., Pan S., Qiu L., S L., and Song Z., “Analytic algorithm of pose estimation based on two pairs of parallel lines,” Chin. J. Sci. Instrum., vol. 29, no. 3, pp. 600, Oct. 2008.
[19] Philips, “LED-luminares,” http://www.lighting.philips.com.cn/prof/led-lamps-and-systems/led-tubes#pfpath=0-LED_GR, 2020. [Online, accessed 8-Apr-2020].
[20] L. Zhou, Y. Yang, M. Abello, and M. Kaess, “A robust and efficient algorithm for the PnL problem using algebraic distance to approximate the reprojection distance,” in Proc. AAAI Conf. Artif. Intell., 2019, vol. 33, pp. 9307–9315.