




Article

Pose Estimation for Visible Light Systems Using a Quadrature Angular Diversity Aperture Receiver

Shengqiang Shen ^{1,2,*} , Jose Miguel Menéndez Sánchez ^{3,4} , Shiyin Li ^{1,2} and Heidi Steendam ³ 

¹ School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221000, China; lishiyin@cumt.edu.cn

² Engineering Research Center of Intelligent Control for Underground Space, Ministry of Education, China University of Mining and Technology, Xuzhou 221000, China

³ Department of Telecommunications and Information Processing, Ghent University/IMEC, Interuniversity Microelectronics Centre, 9000 Ghent, Belgium; jmmenend@espol.edu.ec (J.M.M.S.); heidi.steendam@ugent.be (H.S.)

⁴ Facultad de Ingeniería en Electricidad y Computación, ESPOL Polytechnic University, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, Guayaquil 90112, Ecuador

* Correspondence: sshen@cumt.edu.cn

Abstract: The quadrature angular diversity aperture (QADA) receiver, consisting of a quadrant photodiode (QPD) and an aperture placed above the QPD, has been investigated for pose estimation for visible light systems. Current work on pose estimation for the QADA receiver uses classical camera sensor algorithms well known in computer vision. To this end, however, the light spot center first has to be obtained based on the RSS. However, this is less straightforward than for camera sensors, as in contrast to such sensors where the relationships are linear, the RSS output from the QADA is a non-linear function of the light spot position. When applying closed form solutions or iterative methods for cameras on a QADA, the non-linearity will degrade their performance. Furthermore, since in practice the aperture is not always perfectly aligned with the QPD, a procedure to calibrate the receiver is needed. Current work on calibration requires additional sophisticated equipment to measure the pose during calibration, which increases the difficulty of implementation. In this paper, we target the above problems for pose estimation and calibration of the QADA receiver. To this end, we first study the effect of the strategy of differencing and normalization on the probability density function (PDF), a commonly applied strategy for the QPD's robustness against RSS variation, and it is shown that the applied strategy results in a complex PDF, which makes an effective and efficient estimation hard to achieve. Therefore, we derive an approximated PDF in a simple closed-form, based on which the calibration and the pose estimation algorithms using the least squares principle are proposed. The proposed calibration does not require any information about the pose of the receiver and is robust to variation of the received power and imperfect knowledge of the radiation pattern of the LED, making it easy to implement. We also derive the corresponding Cramér-Rao lower bound on the misalignment to benchmark the performance of the misalignment and to serve as an indicator to determine the required signal-to-noise ratio (SNR) or number of LEDs to obtain a desired accuracy. The calibration and pose estimation are evaluated by means of a Monte Carlo simulation. Computer simulations show that this theoretical bound is close to the RMSE of the proposed estimator and that the proposed pose estimator outperforms the PnP algorithm.

Keywords: pose estimation; calibration; quadrature angular diversity aperture receiver; Cramér-Rao bound; visible light system



Citation: Shen, S.; Menéndez Sánchez, J.M.; Li, S.; Steendam, H. Pose Estimation for Visible Light Systems Using a Quadrature Angular Diversity Aperture Receiver. *Sensors* **2022**, *22*, 5073. <https://doi.org/10.3390/s22145073>

Academic Editor: Mario Muñoz-Organero

Received: 31 March 2022

Accepted: 24 June 2022

Published: 6 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

White LEDs can be modulated up to several MHz. Therefore, they can be used for wireless communication and positioning. As white LEDs are becoming the primary source of light, visible light positioning (VLP) has the potential to achieve positioning with low

cost, low power consumption, and long lifetime [1]. Due to the dependency of the RSS values on the pose (position and orientation) of the receiver, it is expected that imperfect knowledge of the pose may strongly affect the performance of visible light communication (VLC) systems [2–7]. Recently, a compact receiver based on a quadrant PD (QPD) combined with a single aperture, i.e., the quadrature angular diversity aperture receiver (QADA), was proposed for VLP [8]. The QPD is a segmented photodiode consisting of four matched PDs on a single chip separated by very thin gaps. Combined with an aperture, the resulting receiver achieves angular diversity with a reasonably compact size. Furthermore, as only one aperture needs to be placed above the QPD, the receiver will only have a limited number of construction parameters, making it more robust to construction errors.

The light from an LED reaches the QPD through the aperture only and introduces a light spot of the same size and shape as the aperture. As the position of this light spot is determined by the position of the LED and the pose of the receiver, the overlap area between the light spot and the four quadrants contains information about the pose of the receiver. The RSS values in the four constituent PDs of the QPD are proportional to this overlap area, indicating that a comparison of these RSS values allows us to estimate the angle-of-arrival (AOA) for the different LEDs as well as the position and/or orientation (pose) of the receiver. The current work has been initiated to achieve this goal. Experiments in [8] show that high angular resolution in estimated AOA can be achieved. The positioning performance of the QADA receiver is investigated in several works [8–11]. The paper [12] considers 3D position estimation with 1D orientation estimation, i.e., it is able to determine the azimuth angle for a receiver that is pointed upwards. The paper [13] that focuses more on the communication aspects considers the effect of a tilt angle, which restricts the orientation to two degrees of freedom. Recently, the work [14] studied the joint 3D position and 3D orientation estimation, namely pose estimation, based on a classic non-iterative perspective-n-point (PnP) algorithm, a well known algorithm for camera sensors in computer vision. In their work, to employ the PnP, first the light spot center is obtained based on the RSS. However, this is less straightforward than for camera sensors, as in such sensors the relationships are linear, whereas the RSS output from the QADA is a non-linear function of the light spot position. Moreover, the strategy of differencing and normalization is commonly applied to the RSS, claimed by most works to improve the robustness against RSS variation, which further increases the non-linearity. These effects will degrade the performance of the PnP algorithm when applied to the QADA receiver; the PnP algorithm is sensitive to noise and as a result it lacks precision [15].

Furthermore, since in practice the aperture is not always perfectly aligned with the QPD, a procedure to calibrate the receiver is needed. The current work starts to consider the presence of misalignment between the aperture and the QPD. In [8], the authors manually adjusted the position of the aperture as a pre-step in the experiment, and the need for calibration is further emphasized in [9]. However, such a hardware calibration is a burdensome, time-consuming procedure. A better approach is to analyze the effect of the aperture misalignment and compensate for it through signal processing. For example, in [12], the authors propose a calibration method to compensate for the mismatch using signal processing. As the authors assume in this calibration phase that the relative distance between the LED and the receiver is known, additional equipment such as Optitrack is required to determine the ground truth on this distance. This limits the usability of the considered method.

In this paper, we target the above problems for pose estimation and calibration of the QADA receiver. On the one hand, we propose a calibration procedure using signal processing that does not require the knowledge of the pose of the receiver; thus no additional hardware is needed to obtain the ground truth on this pose. On the other hand, a pose estimation algorithm is proposed that directly estimates the pose from the (normalized) RSS values, achieving better performance compared to the method that estimates the light spot center and then applies the camera algorithm. To this end, we first investigate the strategy of differencing and normalization to the RSS, i.e., where we only look at the

relative differences between the normalized RSS values to obtain an observation that is robust against variation in received power and offsets common to all quadrants [12]. We investigate the distribution of the observation obtained from the noisy RSS values, based on which we then propose the algorithm to estimate the pose of the receiver that is more robust to noise.

The main contributions of this paper are as follows:

1. We first model the RSS vector of the QADA at an arbitrary pose using the perspective projection model. With the help of this model, we derive an explicit expression relating the RSS with the intrinsic parameters, i.e., the aperture height and the misalignment of the aperture, and the extrinsic parameters, i.e., the position and orientation of the receiver.
2. We use the strategy of the normalized differences of the RSS values to make the estimator robust against variations in the transmitted power and radiation pattern. To be able to derive a simple estimation algorithm, we replace the complex true PDF of the resulting observations by an approximated Gaussian PDF based on the first order Taylor series approximation of the observation.
3. Using this approximated PDF, we propose a calibration algorithm based on the least squares (LS) principle. The algorithm jointly estimates the intrinsic and extrinsic parameters from which the intrinsic parameters are extracted. This estimation is performed in an iterative way, where the principle of optimization on manifolds is used. After calibration, a simplified version of the estimation algorithm is used to extract the pose, as now we use the calibrated intrinsic parameters as prior knowledge for the misalignment.
4. To evaluate the optimality of the proposed algorithms, we derive the misspecified Cramér-Rao bound (MCRB) to take into account the effect of the approximated PDF. We compare the MCRB with the Cramér-Rao bound (CRB) for the detected RSS values to quantify the performance loss due to using the normalized differences of the RSS values to make the estimator robust against imperfect knowledge of the transmitted power and radiation pattern. The designed algorithms are verified by Monte Carlo simulations.

The rest of the paper is organized as follows. The channel link model and the expressions for the observation and output vectors are presented in Section 2. The approximated PDF and the calibration system model are provided in this section as well. The calibration and the pose estimation algorithms using the LS principle are proposed in Section 3. Subsequently, the theoretical bound is derived in Section 4. The Monte Carlo numerical comparison is given in Section 5. Finally, some concluding remarks are given in Appendix A.

Notation: Scalars are denoted in italic, e.g., x . Lowercase boldface indicates a column vector, e.g., \mathbf{x} . Uppercase boldface denotes a matrix or a set, e.g., \mathbf{X} , with \mathbf{I}_N representing an $N \times N$ identity matrix, and $\mathbf{0}_{N \times M}$ representing an $N \times M$ zero matrix. The vector \mathbf{e}_i is a unit basis vector with its i th element being 1, and the operator $\bar{\cdot}$ converts the Cartesian coordinates \mathbf{x} into the homogeneous coordinates $\bar{\mathbf{x}} = [\mathbf{x}^T, 1]^T$. Matrix transpose and inverse are indicated by superscript T, and -1 , respectively. The Euclidean norm is denoted by $\|\cdot\|$, and the expectation operator is denoted by $E\{\cdot\}$. $\nabla_{\mathbf{x}} = \partial/\partial\mathbf{x}$ denotes the Del operator, while $\Delta_{\mathbf{x}}^z = \nabla_z \nabla_{\mathbf{x}}^T$ denotes the Hessian operator. The rectangular function $\Pi(\cdot)$ is defined as

$$\Pi(x) \triangleq \begin{cases} 1, & |x| \leq 1. \\ 0, & |x| > 1. \end{cases} \quad (1)$$

The set of all real numbers is denoted by \mathbb{R} . The group of all rotation matrices, i.e., the special orthogonal group, is denoted by $SO(3)$, and the associated Lie algebra is denoted

by $\mathfrak{so}(3)$. The special Euclidean group is denoted by $SE(3)$. The operator \cdot^\wedge converts a 6×1 vector into a member of $\mathfrak{se}(3)$ – the Lie algebra of $SE(3)$ – which is,

$$\left(\begin{bmatrix} \mathbf{a}^\top & \mathbf{b}^\top \end{bmatrix}^\top \right)^\wedge = \begin{bmatrix} \mathbf{b}_\times & \mathbf{a} \\ \mathbf{0}^\top & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 4}, \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}^{3 \times 1}, \quad (2)$$

while the operator \cdot_\times converts a 3×1 vector into $\mathfrak{so}(3)$, which is,

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}. \quad (3)$$

The operators \cdot^\vee and \cdot_\otimes are, respectively, the inverse operators of \cdot^\wedge and \cdot_\times . Additionally, $\exp(\cdot)$ and $\log(\cdot)$ are the matrix exponential and matrix logarithm functions, respectively.

2. System Description

In this section, we describe a visible light system that is able to detect signals transmitted by different LEDs from which the pose of the receiver can be determined. To avoid interference between the reference signals for pose estimation, we assume the system adopts multicarrier multiplexing based on (DC-biased) sinusoidal waves [16,17], where a portion of the subcarriers is assigned to the LED anchors to be used as reference signals for pose estimation, while the rest of the subcarriers can be used for communication. Each anchor is assigned a different subcarrier, implying that due to the orthogonality of the subcarriers, the reference signals from the different LEDs can be separated in a simple way. We assume the receiver has knowledge of the used subcarrier and position of each LED, e.g., because this information is contained in the reference signal or is available in a database accessible by the receiver.

The rest of this section is organized as follows. We first model the channel link between a single LED and the QADA and give the expressions for the output vector and observation. We further derive the PDF of the observation. As this PDF is too complex for practical estimation purposes, we also derive a simple closed-form approximation. Finally, we design a system model for calibration and discuss the goals of the pose estimation problem.

2.1. Receiver Structure

The considered visible light system uses a QPD consisting of four matched photodiodes deposited on a single chip separated by very thin gaps. To convert the incident light of an LED into a light spot on the QPD, we place an opaque screen with a circular aperture at a specific height above and parallel to the surface of the QPD [8] as shown in Figure 1 (a square aperture has the disadvantage of being sensitive to rotational offsets, which means that the rotational offset must be calibrated and compensated for, while a circular aperture is circularly symmetric, implying it is inherently robust to rotational offsets. Therefore, in this paper, we focus on the circular aperture). The light spot on the QPD's surface, as shown in Figure 2, leads the QPD to produce a unique quad-tuple of RSS values proportional to the overlap areas, i.e., the RSS output is uniquely determined by the position of the light spot. Because of this, the position of the light spot can be estimated from the RSS values, which in turn, e.g., is useful in detecting an object's angle coordinates [8].

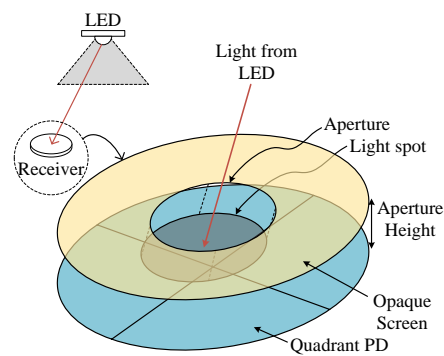


Figure 1. Structure of QADA.

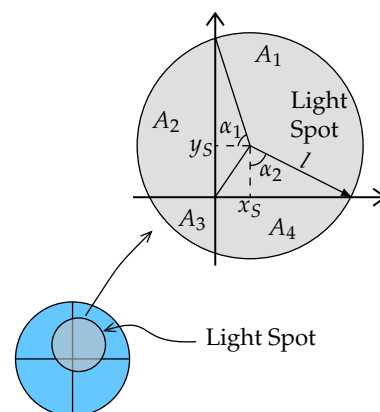


Figure 2. Geometry of the light spot of a light beam, where $x_S = [r_{S,q}]_1$ and $y_S = [r_{S,q}]_2$.

The aperture is chosen so that its diameter is very large compared to the wavelength of the light. Further, as the distance between the LED and the QPD is very large compared to the wavelength, the receiver is in the far-field region of the LED, implying the incident waves can be considered as plane waves. As a result [18], the only light that reaches the QPD is the light that passes through the aperture, and the incident light will introduce a circular light spot of the same size as the aperture on the plane of the QPD. The position of the light spot in the QPD plane not only depends on the position of the light source, but also on the pose of the receiver, implying the RSS outputs in the constituent PDs of the QPD are functions of the position of the light source and the pose of the receiver. We assume that the light spot overlaps with all quadrants and completely falls within the active area of the QPD (in case the light spot does not overlap with all quadrants, one or more of the RSS values will be zero. This will obstruct the strategy of differencing and normalization of the RSS values, which is used to make the estimation algorithm robust against variations in the transmitted optical power. In case the light spot partially falls off the QPD, the relationship between the RSS values and the pose of the receiver becomes very complex, implying the resulting estimation algorithm will have high complexity). To ensure that the first condition is fulfilled, we assume the receiver considers only LEDs that lead all quadrants to produce a sufficiently high RSS output. The second condition implies that the aperture diameter is upper limited by (roughly) the dimension of a quadrant, i.e., half the diameter of a circular QPD or half the side of a square QPD. This guarantees a simple unique relationship between the quad-tuple of RSS values generated by the QPD and the pose of the receiver, and by combining the information from different light sources satisfying the above conditions, the pose of the receiver can easily be determined.

In the following, we will describe the relationship between the pose of the QADA and the position of the light spot, and then we will derive the relationship between the light spot and the RSS output.

2.2. Channel Link Model

To obtain the quad-tuple of RSS values as a function of the pose, we note that the dependency of the RSS values on the pose is twofold. Firstly, the channel gain is determined by the radiation pattern of the LED and the irradiance attenuation, implying the incident power at the QPD (and thus the optical power in the light spot) depends on the relative position of the light spot with respect to the LED. Secondly, the RSS values in the quadrants are proportional to the overlap areas of the light spot with the quadrants of the QPD, which is determined by the position of this light spot with respect to the center of the QPD. To express the RSS values as function of the pose of the receiver, we define the following three coordinate systems: (1) The first coordinate system, i.e., the system frame, corresponds to the inertia frame and is used to describe the pose of the receiver with respect to the LEDs. Furthermore, we define two coordinate systems that are attached to the receiver and will be used to describe the position of the light spot and misalignment due to construction errors; (2) In the first receiver attached coordinate system, i.e., the receiver frame, the x - y plane is the plane of the aperture, and the z -axis is the normal to that plane. The origin O_r of this receiver frame coincides with the center of the aperture. We will use an extra subscript r to indicate coordinates represented in the receiver frame, i.e., the coordinates \mathbf{x}_r in the receiver frame correspond to their counterpart \mathbf{x} in the system frame. This coordinate system is used to find the distance vector of the light spot to the LED as a function of the pose; (3) The second receiver attached coordinate system is the QPD frame, where the x - y plane is the surface of the QPD, the origin O_q is placed in the center of the QPD, and the x and y axes are aligned with the gaps between the quadrants. We will use an extra subscript q to indicate coordinates represented in the QPD frame, i.e., the coordinates \mathbf{x}_q in the QPD frame correspond to their counterpart \mathbf{x} in the system frame. Although this coordinate system is closely related to the receiver frame, its benefit lies in the simplification of the computations, as it enables us to convert the 3D receiver frame coordinates to the 2D coordinates in the QPD frame (we assume that the aperture plane is parallel to the plane of the QPD, and the x - and y -axes of the QPD and receiver frames are aligned. This implies that the transformation of vectors in the aperture frame into vectors in the QPD frame is a simple translation). This approach is commonly used in the perspective projection model, which we employ to determine the position of the light spot in the QPD plane, and from which we compute the overlap areas of the light spot and the QPD. In the following, we will determine the coordinates of the light spot generated by an LED.

Assume that in the system frame, the pose of the receiver is described by its orientation $\mathbf{R} \in SO(3)$, i.e., \mathbf{R} belongs to the Special Orthogonal group and position $\mathbf{r} \in \mathbb{R}^{3 \times 1}$, where the position of the receiver is defined as the position of the center of the aperture. We further assume the LED has position \mathbf{r}_L and normal vector \mathbf{n}_L in the system frame. To calculate the position of the light spot in the QPD frame, we notice that the line going through the position \mathbf{r}_L of the LED and the centers \mathbf{r} of the aperture and $\mathbf{r}_S \in \mathbb{R}^{3 \times 1}$ of the light spot can be seen as a ray in the perspective projection model of a center of projection located at the center of the aperture and having focal length equal to the aperture height $|h|$, i.e., the distance between the planes of the aperture and the QPD (see Figure 3). The perspective projection model allows us to relate the 3D coordinates of the LED—the ‘object’—in the system frame to the 2D coordinates of the light spot—the ‘image’—in the QPD frame. Note that the position of the light spot in the QPD frame is specified by the (x, y) coordinates only. To make the paper self-contained, we include the background on the perspective projection model. Taking this into account and using the projection principle, the position $\mathbf{r}_{S,q} \in \mathbb{R}^{2 \times 1}$ of the center of the light spot in the QPD frame is given by [19]:

$$\mathbf{r}_{S,q} = \mathbf{k}_{m,q} + \underbrace{\begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{bmatrix}}_B \underbrace{[\mathbf{I}_3, \mathbf{0}] \frac{h}{\mathbf{e}_3^T \mathbf{R} \mathbf{r}_L}}_C \underbrace{\mathbf{R} \mathbf{r}_L}_A. \quad (4)$$

The matrix $\mathbf{T} \in SE(3)$ is defined through its inverse

$$\mathbf{T}^{-1} = \begin{bmatrix} \mathbf{R} & \mathbf{r} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}. \quad (5)$$

Furthermore, the misalignment vector $\mathbf{k}_{m,q}$ expresses the errors in the alignment of the aperture with respect to the QPD. Due to construction errors, a small horizontal offset between the origins of the QPD frame and the receiver frame may occur. This is illustrated in Figure 3, where the brown dot does not coincide with the center O_q of the QPD. The horizontal misalignment vector, which takes into account this offset in the transformation of the light spot coordinates in the receiver frame to the coordinates in the QPD frame, is defined as $\mathbf{k}_{m,q} \doteq [u_m, v_m]^T$, with u_m and v_m the misalignment in the x and y direction, respectively. Taking into account (4), we see that the projection principle consists of three steps:

- Step A: This step converts the LED coordinates $\bar{\mathbf{r}}_L$ in the system frame to the coordinates $\bar{\mathbf{r}}_{L,r} = \mathbf{T}\bar{\mathbf{r}}_L$ in the receiver frame using the matrix \mathbf{T} .
- Step B: This scales $\bar{\mathbf{r}}_{L,r}$ along the projection line with the factor $\frac{h}{h_L}$ where $h_L = \mathbf{e}_3^T \mathbf{T}\bar{\mathbf{r}}_L$ and gets rid of the last component of the homogeneous coordinates through the 3×4 matrix $[\mathbf{I}_3, \mathbf{0}]$, consisting of the 3×3 identity matrix extended with a zero column, as this last component is irrelevant for the determination of the position of the light spot. This results in the position $\mathbf{r}_{S,r} = [\mathbf{I}_3, \mathbf{0}] \frac{h}{\mathbf{e}_3^T \mathbf{T}\bar{\mathbf{r}}_L} \bar{\mathbf{r}}_{L,r}$ of the light spot in the receiver frame.
- Step C: This transforms $\mathbf{r}_{S,r}$ into the 2D coordinates $\mathbf{r}_{S,q}$ by discarding the z coordinate through the mapping of $\mathbf{r}_{S,r}$ on the 2×3 matrix $\begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{bmatrix}$ and adds the misalignment vector $\mathbf{k}_{m,q}$.

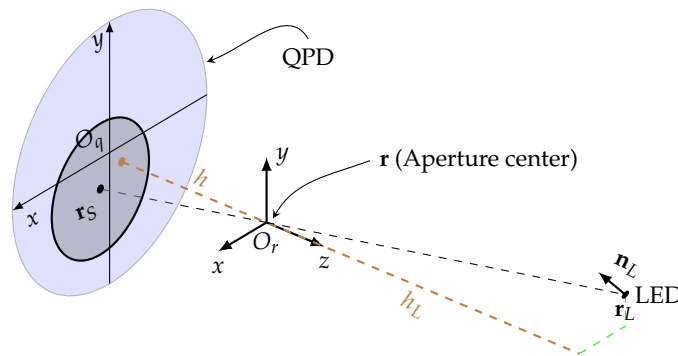


Figure 3. The channel link model. The position of the LED and centers of the aperture and the light spot comply with the perspective projection model, while the shape and size of the light spot is determined by the aperture. Note that in this perspective projection model, the light spot does not correspond to the image of the LED, but it only relates the position of the light spot with the position of the LED.

Equation (4) can be rewritten in the following compact form:

$$\mathbf{r}_{S,q} = \frac{\mathbf{K}\mathbf{T}\bar{\mathbf{r}}_L}{\mathbf{e}_3^T \mathbf{T}\bar{\mathbf{r}}_L}. \quad (6)$$

The matrix \mathbf{K} , which is used in the perspective projection model, is called the intrinsic parameter matrix and describes the internal properties of a receiver, i.e., that are independent of the pose. It only depends on the misalignment vector $\mathbf{k}_{m,q}$ and the aperture height h :

$$\mathbf{K} = \begin{bmatrix} h & 0 & u_m & 0 \\ 0 & h & v_m & 0 \end{bmatrix}. \quad (7)$$

On the other hand, the matrix \mathbf{T} contains the pose information of the receiver and is independent of the internal properties of the receiver. Therefore, it is called the extrinsic parameter matrix. Similarly, taking into account (4), we can write the position $\mathbf{r}_{S,r}$ of the center of the light spot in the receiver frame (in homogeneous coordinates) in the following compact form:

$$\bar{\mathbf{r}}_{S,r} = \frac{\mathbf{K}_0 \mathbf{T} \bar{\mathbf{r}}_L}{\mathbf{e}_3^T \mathbf{T} \bar{\mathbf{r}}_L} \quad \text{with} \quad \mathbf{K}_0 = \begin{bmatrix} h \mathbf{I}_3 & \mathbf{0} \\ \mathbf{e}_3^T & 0 \end{bmatrix}, \quad (8)$$

and the position $\bar{\mathbf{r}}_S$ of the light spot in the system frame (in homogeneous coordinates) is obtained through:

$$\bar{\mathbf{r}}_S = \mathbf{T}^{-1} \bar{\mathbf{r}}_{S,r}. \quad (9)$$

To determine the RSS values in the quadrants of the QPD, we first need to compute the overlap area between the light spot and the quadrants of the QPD. Let us assume the aperture has radius l . Taking into account we already assumed that the light spot completely overlaps with the QPD, i.e., no part of the light spot falls outside the QPD, and that all quadrants overlap with the light spot, the overlap areas can easily be computed based on the position $\mathbf{r}_{S,q}$ of the center of the light spot in the frame of the QPD. Defining the overlap area $A_i > 0$ of the i th quadrant, $i = 1, \dots, 4$ and the angles $\alpha_j \in [0, \pi]$ with $[\mathbf{r}_{S,q}]_j = l \cos \alpha_j$, $j = 1, 2$, and the operator $[\mathbf{r}_{S,q}]_j$ as the j th element of $\mathbf{r}_{S,q}$ (see Figure 2), we obtain the overlap areas by solving the following set of equations:

- The light spot has area πl^2 , implying $A_1 + A_2 + A_3 + A_4 = \pi l^2$.
- Taking into account that the area of a circular segment with central angle 2α , $\alpha \in [0, \pi]$ equals $\frac{l^2}{2}(2\alpha - \sin 2\alpha)$, we find that $A_2 + A_3 = (\alpha_1 - \sin \alpha_1 \cos \alpha_1) l^2$ and $A_3 + A_4 = (\alpha_2 - \sin \alpha_2 \cos \alpha_2) l^2$.
- The overlap area A_1 with the first quadrant is a combination of a circle sector with central angle $\frac{3\pi}{2} - (\alpha_1 + \alpha_2)$ having an area $(\frac{3\pi}{2} - (\alpha_1 + \alpha_2)) \frac{l^2}{2}$, a rectangle with area $[\mathbf{r}_{S,q}]_1 [\mathbf{r}_{S,q}]_2 = l^2 \cos \alpha_1 \cos \alpha_2$ and two triangles with areas $\frac{l^2}{2} \cos \alpha_j \sin \alpha_j$, $j \in \{1, 2\}$.

Note that, depending whether $\cos \alpha_j > 0$ (if $\alpha_j \in [0, \frac{\pi}{2}]$) or $\cos \alpha_j < 0$ (if $\alpha_j \in [\frac{\pi}{2}, \pi]$), the above contributions will combine positively or negatively, resulting in the wanted overlap area A_i . Hence, the overlap area vector $\mathbf{A} = [A_1, \dots, A_4]^T$ is the solution of $\mathbf{M}_a \mathbf{A} = \mathbf{s}$ with

$$\mathbf{M}_a = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{s} = l^2 \begin{bmatrix} \frac{3\pi}{4} - \frac{\alpha_1 + \alpha_2}{2} + \frac{\sin 2\alpha_1 + \sin 2\alpha_2}{4} + \cos \alpha_1 \cos \alpha_2 \\ (\alpha_1 - \sin \alpha_1 \cos \alpha_1) \\ (\alpha_2 - \sin \alpha_2 \cos \alpha_2) \\ \pi \end{bmatrix}. \quad (10)$$

Now that we obtained the overlap areas of the light spot with the PDs of the QPD and related them to the pose of the receiver, the RSS values in the different PDs can be determined. To this end, we define the distance vector between the transmitter and the light spot as $\mathbf{v} = \mathbf{r}_S - \mathbf{r}_L$. With this definition, we can write the distance v between the LED and the light spot, and the incidence angle θ , i.e., the angle between the vector $-\mathbf{v}$ and the normal $\mathbf{n}_Q = \mathbf{R} \mathbf{e}_3$ of the QPD as:

$$v = \|\mathbf{v}\|, \quad (11)$$

$$\cos(\theta) = -\frac{(\mathbf{T}^{-1} \mathbf{e}_3)^T \bar{\mathbf{v}}}{\|\bar{\mathbf{v}}\|}, \quad (12)$$

where the last equation holds due to the equalities $[\mathbf{n}_Q^T, 0]^T = \mathbf{T}^{-1} \mathbf{e}_3$, obtained from (5), and $\mathbf{n}_Q^T \mathbf{v} = [\mathbf{n}_Q^T, 0]^T \bar{\mathbf{v}}$. The channel gain h_i for the i th quadrant of the QPD is given by [20]

$$h_i = \frac{\Gamma(\mathbf{n}_L, \mathbf{v}) \cos \theta \Pi_\theta}{2\pi v^2} A_i, \quad (13)$$

where $\Gamma(\mathbf{n}_L, \mathbf{v})$ is the radiation pattern of the LED pointing \mathbf{n}_L and is evaluated at direction \mathbf{v} , $\Pi_\theta = \Pi(\phi/\phi_{FOV})$. The factor Π_θ in (13) implies that the QPD can detect the light only when the LED is within its FOV, i.e., $0 \leq \phi \leq \phi_{FOV}$. Taking into account (11)–(13), the channel gain vector is given by

$$\mathbf{h}(\mathbf{k}, \mathbf{T}) = -\frac{\Pi_\theta \Gamma(\mathbf{n}_L, \mathbf{v})(\mathbf{T}^{-1} \mathbf{e}_3)^T \bar{\mathbf{v}}}{2\pi \|\mathbf{v}\|^3} \mathbf{A}, \quad (14)$$

and is a function of the intrinsic parameter vector $\mathbf{k} \doteq [u_m, v_m, h]^T$ and the pose \mathbf{T} . From this channel gain, we obtain the RSS output vector $\mathbf{y} = [y_1, \dots, y_4]^T$ in the four PDs:

$$\mathbf{y} = \mathbf{g} + \boldsymbol{\eta}, \quad (15)$$

where $\mathbf{g} = R_p P_t \mathbf{h}$, with R_p the responsivity of the QPD, and P_t the transmitted power. Due to coupling between the quadrants, the noise components are correlated Gaussian distributed random variables $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma_\eta^2 \mathbf{C})$, where \mathbf{C} is the correlation matrix with $C_{i,i} = 1$ and $C_{i,j} = \rho$ if $i \neq j$, ρ is the correlation between the noise of different quadrants, which can be determined by comparing the noise values in the different quadrants during an experiment and calculate the resulting correlation, and σ_η^2 is the noise variance.

In the derivation of \mathbf{g} in (15), it is assumed that the transmitted optical power is perfectly known, and the LED is a perfect Lambertian radiator. However, directly estimating the intrinsic and/or extrinsic parameters from the RSS output vector will result in inaccurate estimates. To make the estimator more robust to changes in the incident power, the normalized differences between the RSS values along the two axes can be used, as

$$\begin{cases} t_x = \frac{(y_1+y_4)-(y_2+y_3)}{y_1+y_2+y_3+y_4} \\ t_y = \frac{(y_1+y_2)-(y_3+y_4)}{y_1+y_2+y_3+y_4} \end{cases}. \quad (16)$$

The approach (16) will make the observations insensitive to variations in the transmitted optical power and radiation pattern provided that the RSS values are noise-free or affected by a common offset so that the normalized differences of (16) are noiseless observations. However, the noise induced by ambient light in visible light systems is non-negligible, and due to the randomness of the shot and thermal noise, the noise present in the different quadrants is in general not equal. The presence of this non-identical random noise will have an impact on the distribution of t_x and t_y . As the estimation of the intrinsic and extrinsic parameters relies on proper knowledge of this distribution, we need to evaluate the effect of the noise on the distribution to see if the observations (16) are still insensitive to variations in the transmitted power when noise is present. Therefore, in the following section, we will analyze the distribution of the normalized differences t_x and t_y .

2.3. Approximation to the PDF of the Observation

To find the distribution of the normalized differences, we first rewrite the numerator and denominator of t_x and t_y in a more compact form by introducing $\mathbf{m} \doteq [m_1, m_2, m_3]^T = \mathbf{M}_y \mathbf{g}$ and $\mathbf{w} \doteq [w_1, w_2, w_3]^T = \mathbf{M}_y \boldsymbol{\eta}$ with

$$\mathbf{M}_y = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad (17)$$

implying (16) can be rewritten as $\mathbf{t} = \left[\frac{m_1+w_1}{m_3+w_3}, \frac{m_2+w_2}{m_3+w_3} \right]^T$. Taking into account that the components of \mathbf{w} are linear combinations of Gaussian distributed random variables, \mathbf{w} has a multivariate normal distribution, i.e., $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_w)$ with

$$\boldsymbol{\Sigma}_w = 4\sigma_\eta^2 \begin{bmatrix} 1-\rho & 0 & 0 \\ 0 & 1-\rho & 0 \\ 0 & 0 & 1+3\rho \end{bmatrix}. \quad (18)$$

Note that the numerator $(m_3 + w_3) = \sum_i y_i$ corresponds to the sum of the RSS values in the different quadrants. Hence, m_3 and $\sigma_{w_3}^2 \doteq 4(1 + 3\rho)\sigma_{\eta}^2$ are the average and variance of the total received signal strength, which implies we can define $\zeta \doteq m_3^2/\sigma_{w_3}^2$ as the received signal-to-noise ratio (SNR) at the receiver. Furthermore, we define

$$\boldsymbol{\mu} \doteq \begin{bmatrix} \frac{m_1}{m_3} & \frac{m_2}{m_3} \end{bmatrix}^T, \quad (19)$$

i.e., the ratio of the averages of the denominator and numerator of \mathbf{t} . Obviously, in the absence of noise, the observation equals $\mathbf{t} = \boldsymbol{\mu}$. We will show that:

- $\boldsymbol{\mu}$ is independent of the channel gain parameters and the transmitted power.
- At high SNR, the distribution of \mathbf{t} is given by $\mathbf{t} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\Sigma} = \zeta^{-1} \left(\frac{1-\rho}{1+3\rho} \mathbf{I}_2 + \boldsymbol{\mu} \cdot \boldsymbol{\mu}^T \right). \quad (20)$$

Let us first take a closer look at the dependency of $\boldsymbol{\mu}$ on the channel gain and the transmitted power. Defining $\mathbf{h}(\mathbf{k}, \mathbf{T}) = \lambda \mathbf{A}$, and taking into account that the factor λ is independent of the considered quadrant, it follows that $\boldsymbol{\mu}$ is independent of the channel gain parameters and the transmitted power contained in the factor λ , but only depends on the overlap area vector \mathbf{A} . Substituting the solution of $\mathbf{M}_a \mathbf{A} = \mathbf{s}$ in (19), we find after some straightforward computations that

$$[\boldsymbol{\mu}]_i = \left(1 - \frac{2(\alpha_i - \cos \alpha_i \sin \alpha_i)}{\pi} \right). \quad (21)$$

Hence, the vector $\boldsymbol{\mu}$ solely depends on the angles α_i that are determined by the coordinates of the center of the light spot in the QPD frame. These coordinates, and therefore also $\boldsymbol{\mu}$, only depend on the intrinsic and extrinsic parameters through (6) but not on the channel gain parameters nor the transmitted power.

To obtain the distribution of \mathbf{t} at high SNR, we first notice that $t_x = \frac{m_1+w_1}{m_3+w_3}$ and $t_y = \frac{m_2+w_2}{m_3+w_3}$ are both a ratio of two independent normal variables with non-zero means, and as they share the same denominator, they are correlated. The distribution of a single ratio of normal variables has been investigated in the literature, e.g., [21–23]. In these works, it is shown that a closed-form expression for the PDF in the general case for such a ratio is very complex as its shape differs significantly with its parameters, i.e., in some cases it resembles the Cauchy distribution, in other cases a normal distribution or a bimodal distribution. Several closed-form approximations were discussed in the literature, e.g., the approximate normal distribution. In this paper, we need the distribution of two correlated ratios of normal variables. As this is an even more complicated situation compared to the single ratio case, it is clear that the closed-form expression for the PDF will be even more complex. To simplify the analysis, we first notice that at high SNR, $\frac{w_i}{m_3} \ll 1$. Expanding \mathbf{t} with respect to \mathbf{w} using the Taylor series, keeping up to the linear terms in \mathbf{w} , we obtain:

$$\mathbf{t} \approx \boldsymbol{\mu} + \mathbf{M}_w \mathbf{w} \quad \text{with} \quad \mathbf{M}_w = \begin{bmatrix} \frac{1}{m_3} & 0 & -\frac{m_1}{m_3^2} \\ 0 & \frac{1}{m_3} & -\frac{m_2}{m_3^2} \end{bmatrix}. \quad (22)$$

This first-order approximation directly leads to the distribution $\mathbf{t} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = \mathbf{M}_w \boldsymbol{\Sigma}_w \mathbf{M}_w^T$. Substituting \mathbf{M}_w and $\boldsymbol{\Sigma}_w$ (18) in $\boldsymbol{\Sigma}$, and using the definition of the SNR ζ , it follows that the covariance matrix $\boldsymbol{\Sigma}$ reduces to (20).

Following the above analysis, at high SNR we can design an estimator based on $\boldsymbol{\mu}$ to achieve robustness against variations in the transmitted optical power and radiation pattern to estimate parameters from \mathbf{t} .

3. Calibration and Pose Estimation

Using the approximate PDF of the normalized differences derived in the Section 2, we can now estimate the pose of the receiver, taking into account the relations (19) and (21). When the intrinsic parameters are known, the estimation of the pose, i.e., the extrinsic parameters, is straightforward. However, in practice, the intrinsic parameters are not prior known and need to be estimated because inaccurate knowledge of the misalignment parameters may result in significant biases in the pose estimates. As these intrinsic parameters can be considered fixed once the receiver is assembled, the intrinsic parameters can be determined once during a calibration process. The standard approach is to determine the intrinsic parameters assuming the pose of the receiver is known. However, this approach requires accurate knowledge of the pose of the receiver. Although it is possible to accurately determine the pose of the receiver, this requires costly equipment, in particular to determine the orientation of the receiver, and a laborious procedure to determine the pose of the receiver before each measurement. Therefore, we propose in this section a method to estimate the intrinsic parameters not requiring the knowledge of the pose, making it simpler and less costly than the standard calibration procedure. Note that our approach differs from the classical algorithm commonly used in camera-based systems, e.g., the non-iterative PnP algorithm and its iterative version [24], as in our method we directly estimate the pose from the normalized differences, while in the classical algorithm used for cameras, first the position of the light spot is determined, after which the pose is extracted. Although the algorithm used for cameras can be adapted to solve the problem at hand, we expect that this algorithm will perform worse than the algorithm proposed in this paper, as compared to the high resolution pixel measurement from the camera sensor, the RSS output from the QADA typically has a lower SNR. Furthermore, because of the non-linear relationship between the center of the light spot and the RSS values, the algorithm will be suboptimal for the considered system [15,25]. As the proposed method directly estimates the pose from the RSS values, it also differs from classical techniques used to determine the position, i.e., where first the AOA or distances to the LEDs are estimated before determining the position using, e.g., trilateration or triangulation.

3.1. Calibration Procedure

The calibration procedure proposed in this paper makes use of the setup shown in Figure 4, where N_L LEDs are installed on a plane in the system frame $xO_s y$, and the positions of these LEDs are fixed and known. The receiver observes these LEDs at N_T different poses. Hence, the receiver observes for each LED N_T quad-tuples of RSS values. The resulting observations are arranged in the $4N_T N_L \times 1$ vector $\check{\mathbf{y}} \in \mathbb{R}^{4N_T N_L \times 1}$:

$$\check{\mathbf{y}} = \check{\mathbf{g}} + \check{\boldsymbol{\eta}}, \quad (23)$$

where the vector $\check{\boldsymbol{\eta}} \sim \mathcal{N}(\mathbf{0}, \check{\boldsymbol{\Sigma}}_{\boldsymbol{\eta}})$ collects the noise, and $\check{\boldsymbol{\Sigma}}_{\boldsymbol{\eta}} = \sigma_{\boldsymbol{\eta}}^2 \mathbf{I}_{N_T N_L} \otimes \mathbf{C}$ (see Section 2.2 for the definition of the noise correlation matrix \mathbf{C}). The vector $\check{\mathbf{g}}$ can be rewritten as $\check{\mathbf{g}} = R_p \left[P_{t,1} \mathbf{g}_{1,1}^T, P_{t,2} \mathbf{g}_{1,2}^T, \dots, P_{t,N_L} \mathbf{g}_{N_T, N_L}^T \right]^T$, where $\mathbf{g}_{i,j} \in \mathbb{R}^{4 \times 1}$ is the channel gain between the j th LED and the receiver at the i th pose \mathbf{T}_i^{-1} , and $P_{t,j}$ is the transmitted power of the j th LED. For each LED and each pose, we compute the normalized differences and arrange them in the $2N_T N_L \times 1$ vector $\check{\mathbf{t}} \in \mathbb{R}^{2N_T N_L \times 1}$. Taking into account that the noise in the different observations is statistically independent, we can write the Gaussian approximation of the PDF of $\check{\mathbf{t}}$ as:

$$p_m(\check{\mathbf{t}}) = \frac{1}{\sqrt{\det(2\pi\check{\boldsymbol{\Sigma}})}} e^{(-\frac{1}{2}\|\check{\mathbf{t}}-\check{\boldsymbol{\mu}}\|_{\check{\boldsymbol{\Sigma}}}^2)} \quad (24)$$

where $\check{\boldsymbol{\mu}} = \left[\boldsymbol{\mu}_{1,1}^T, \boldsymbol{\mu}_{1,2}^T, \dots, \boldsymbol{\mu}_{N_T, N_L}^T \right]^T$. The components $\boldsymbol{\mu}_{i,j}$ correspond to the j th LED and the i th pose and are defined through (19), while the covariance matrix $\check{\boldsymbol{\Sigma}}$ is the block diagonal matrix given by

$$\check{\Sigma} = \text{diag}(\Sigma_{1,1}, \Sigma_{1,2}, \dots, \Sigma_{1,N_L}, \Sigma_{2,1}, \dots, \Sigma_{N_T, N_L}) \quad (25)$$

In this covariance matrix, $\Sigma_{i,j}$ is the covariance matrix for the noise of $\mathbf{t}_{i,j}$ (see (20)). These covariance matrices $\Sigma_{i,j}$ are a function of $\mu_{i,j}$, which in turn depend on the parameters to be estimated. As a consequence, maximum likelihood estimation will be complex. Therefore, we will use the least squares (LS) method to estimate the intrinsic and extrinsic parameters. Let us define the parameter set $\Theta = \{\mathbf{k}, \mathbf{T}_1, \dots, \mathbf{T}_{N_T}\}$, then the LS estimate is given by:

$$\begin{aligned} \hat{\Theta} &= \arg \min_{\Theta} \frac{1}{2} \|\check{\mathbf{t}} - \check{\boldsymbol{\mu}}\|^2, \\ \text{s.t. } &\mathbf{R}_i^T \mathbf{R}_i = \mathbf{I}, \quad \det(\mathbf{R}_i) = +1. \end{aligned} \quad (26)$$

The constraints $\mathbf{R}_i^T \mathbf{R}_i = \mathbf{I}$ and $\det(\mathbf{R}_i) = +1$ imply that \mathbf{R}_i , which defines the orientation of the receiver at pose i and is enclosed in the transform matrix $\mathbf{T}_i^{-1} \in SE(3)$ (see (5)), is a rotation matrix that is a member of $SO(3)$. Unfortunately, (26) has no closed-form solution, implying we need to resort to an iterative algorithm to estimate Θ . As the solution of the above (iterative) constrained optimization problem is complex and cumbersome in Euclidean space, we first notice that $SE(3)$ is a manifold and convert the above optimization problem into an unconstrained optimization problem on manifolds [26]. In what follows, we use the Gauss–Newton method on manifold $SE(3)$ to solve (26), similarly as in [27]. At each iteration, the update direction is calculated by

$$[\Delta_{\mathbf{k}}, \Delta_{\mathbf{T}_1}, \dots, \Delta_{\mathbf{T}_{N_T}}]^T = (\nabla_{\Theta} \check{\boldsymbol{\mu}})^{\dagger} (\check{\mathbf{t}} - \check{\boldsymbol{\mu}}) \quad (27)$$

where $(\cdot)^{\dagger}$ is the Moore–Penrose pseudoinverse, and $\nabla_{\Theta} \check{\boldsymbol{\mu}} = [\nabla_{\Theta} \mu_{1,1}, \nabla_{\Theta} \mu_{1,2}, \dots, \nabla_{\Theta} \mu_{N_T, N_L}]^T \in \mathbb{R}^{2N_T N_L \times (6N_T + 3)}$ the gradient of $\check{\boldsymbol{\mu}}$ with respect to Θ . The derivation of the components $\nabla_{\Theta} \mu_{i,j}$ can be performed in a similar way as in [27], and the result is given in Appendix A. As the update directions (27) depend on \mathbf{k} and all \mathbf{T}_i , the intrinsic parameters and the poses need to be estimated jointly:

$$\begin{cases} \mathbf{k}^{t+1} = \mathbf{k}^t + \tau_{\mathbf{k}} \Delta_{\mathbf{k}}(\mathbf{k}^t, \mathbf{T}_1^t, \dots, \mathbf{T}_{N_T}^t) \\ \mathbf{T}_i^{t+1} = \exp\left(\left(\tau_{\mathbf{T}} \Delta_{\mathbf{T}_i}(\mathbf{k}^t, \mathbf{T}_1^t, \dots, \mathbf{T}_{N_T}^t)\right)^{\wedge}\right) \mathbf{T}_i^t, \quad i \in \{1, \dots, N_T\} \end{cases} \quad (28)$$

where $\tau_{\mathbf{k}}$ and $\tau_{\mathbf{T}}$ control the incremental step size for \mathbf{k} and \mathbf{T}_i , respectively.

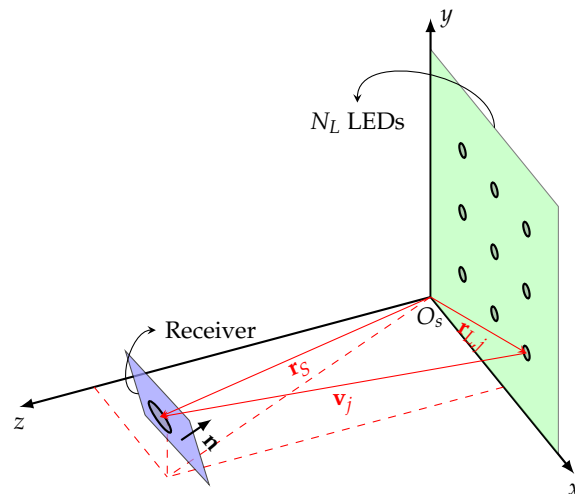


Figure 4. System frame for calibration.

The Gauss–Newton method needs an initialization Θ^0 to start the iterative estimation process. To obtain this initial estimate, we consider the direct linear transformation

method [24]. Assuming the positions of at least four LEDs, and the positions of the light spots for these LEDs are known, the direct linear transformation method gives a closed-form coarse estimate for Θ^0 . Although in our problem, the positions of the light spots are not known, we can estimate them from the normalized differences $\mathbf{t}_{i,j}$, which is a noisy version of $\boldsymbol{\mu}_{i,j}$ (22). This $\boldsymbol{\mu}_{i,j}$ is a function of the angles α_ℓ , $\ell = 1, 2$ (21), which in turn are related to the position $\mathbf{r}_{S,q,i,j}$ of the light spot, i.e., we can write $\boldsymbol{\mu}_{i,j}$ as a function of $\mathbf{r}_{S,q,i,j}$: $\boldsymbol{\mu}_{i,j}(\mathbf{r}_{S,q,i,j})$. In many situations, the noise is relatively small, implying we can neglect the presence of the noise, i.e., $\mathbf{t}_{i,j} \approx \boldsymbol{\mu}_{i,j}$. Inverting the function $\boldsymbol{\mu}_{i,j}(\mathbf{r}_{S,q,i,j})$, we obtain the coarse estimate $\hat{\mathbf{r}}_{S,q,i,j} \in \mathbb{R}^{2 \times 1}$ of the light spot for the j th LED and the i th pose:

$$\hat{\mathbf{r}}_{S,q,i,j} = (\boldsymbol{\mu}_{i,j})^{(-1)}(\mathbf{t}_{i,j}), \quad (29)$$

where $(\cdot)^{(-1)}$ is the inversion operator. Unfortunately, due to the non-linearity of $\boldsymbol{\mu}$ as a function of its argument $\mathbf{r}_{S,q}$, no closed-form expression can be found for this inverse. However, as $\boldsymbol{\mu}$ monotonically increases with $\mathbf{r}_{S,q}$, we can precompute the function $\boldsymbol{\mu}$ for a set of values $\mathbf{r}_{S,q}$ and save them in a look-up table. Based on this table, the inverse can be determined by interpolating between the values available in the look-up table.

3.2. Theoretical Lower Bound

In the Section 3.1, we considered the LS estimation of the intrinsic parameters based on the approximated PDF of $\check{\mathbf{t}}$ in the presence of the (unwanted) extrinsic parameters \mathbf{T}_i . To gain insight into the optimality of the designed calibration algorithm and to investigate how the SNR, N_L or N_T would effect the accuracy of the estimator, in this section we derive the theoretical lower bound for the mean squared errors of the intrinsic parameters h and $\mathbf{k}_{m,q}$. Since the estimation includes the unwanted parameter \mathbf{T}_i and since the estimator is designed based on an approximated PDF, we will derive the misspecified Cramér-Rao bound (MCRB) [28] for the whole parameter set Θ and only keep the left upper 3×3 submatrix corresponding to the intrinsic parameters.

The MCRB for the whole parameter set Θ is given by

$$\text{MCRB}(\Theta) = (\mathbf{M}_1(\hat{\Theta}))^{-1} \mathbf{M}_2(\hat{\Theta}) (\mathbf{M}_1(\hat{\Theta}))^{-1} + \text{Bias}(\hat{\Theta}, \Theta). \quad (30)$$

In (30), $\hat{\Theta}$ is the parameter set to which the estimation converges due to the PDF approximation assumed by the estimator, given by

$$\hat{\Theta} = \arg \min_{\Theta} D(p(\check{\mathbf{t}}) || p_m(\check{\mathbf{t}}|\Theta)), \quad (31)$$

where $D(p(\check{\mathbf{t}}) || p_m(\check{\mathbf{t}}|\Theta))$ is the Kullback–Leibler divergence (KLD) between the true $(p(\check{\mathbf{t}}))$ and the approximated $(p_m(\check{\mathbf{t}}|\Theta))$ PDFs. The matrices \mathbf{M}_1 , \mathbf{M}_2 , and $\text{Bias}(\hat{\Theta}, \Theta)$ are, respectively, given by [28]

$$\mathbf{M}_1 = E_{\check{\mathbf{t}}} \left\{ \Delta_{\Theta}^{\Theta} \ln p_m(\check{\mathbf{t}}|\Theta), \right\} \quad (32)$$

$$\mathbf{M}_2 = E_{\check{\mathbf{t}}} \left\{ \nabla_{\Theta}^T \ln p_m(\check{\mathbf{t}}|\Theta) \nabla_{\Theta} \ln p_m(\check{\mathbf{t}}|\Theta), \right\} \quad (33)$$

$$\text{Bias}(\hat{\Theta}, \Theta) = \boldsymbol{\epsilon}(\hat{\Theta}, \Theta) \boldsymbol{\epsilon}(\hat{\Theta}, \Theta)^T, \quad (34)$$

where $\boldsymbol{\epsilon}(\hat{\Theta}, \Theta)$ is the error vector between $\hat{\Theta}$ and Θ . The error vector is defined as $\boldsymbol{\epsilon}(\hat{\Theta}, \Theta) = [\boldsymbol{\epsilon}_{\mathbf{k}}, \boldsymbol{\epsilon}_{\mathbf{T}_1}, \dots, \boldsymbol{\epsilon}_{\mathbf{T}_{N_T}}]^T$ with $\boldsymbol{\epsilon}_{\mathbf{k}} = \hat{\mathbf{k}} - \mathbf{k}$ and $\boldsymbol{\epsilon}_{\mathbf{T}_i} = \log(\mathbf{T}_i \hat{\mathbf{T}}_i^{-1})^{\vee}$. After a few rearrangements, it can be seen that the expectations in (32) and (33) depend on $p(\check{\mathbf{t}})$ through its mean and covariance only. Due to the complexity of the true PDF $p(\check{\mathbf{t}})$, this mean and average must be calculated numerically via Monte Carlo integration.

To analyze the effect of N_T , N_L and SNR, we take in account that since the approximated PDF $p_m(\check{\mathbf{t}}|\Theta)$ is Gaussian, the minimum of the KLD (31) is obtained when the first two moments of the true and the approximated PDFs are matched [29]. When the SNR is

high, the first two moments of the approximated PDF approach those of the true PDF. As a result, the estimator is asymptotically unbiased, i.e., $\text{Bias}(\hat{\Theta}, \Theta) \approx \mathbf{0}$ for $\text{SNR} \gg 1$, so that $\mathbf{M}_1 \approx \mathbf{M}_2$, $\text{MCRB}(\Theta) \approx \mathbf{M}_2^{-1}$ and

$$\mathbf{M}_2 \approx \sum_i^{N_T} \sum_j^{N_L} \nabla_{\Theta}^T \mu_{i,j} \Sigma_{i,j}^{-1} \nabla_{\Theta} \mu_{i,j}. \quad (35)$$

It can be seen that when N_T or N_L increases, the number of terms in (35) increases. As the matrices being summed in (35) are symmetric and positive-semidefinite, this implies that \mathbf{M}_2 is non-decreasing in the sense of the Loewner order. Furthermore, as $\Sigma_{i,j}^{-1}$ is proportional to the SNR, increasing the SNR also enlarges the Loewner order of these matrices being summed and thus leads the partial order of \mathbf{M}_2 to enhance. As a result, the MCRB is a non-increasing function of N_T , N_L , or SNR. While an increase of N_T , N_L , or SNR improves the robustness of the estimator against noise, these parameters do not have an impact on the feasibility of the estimation. The number of observations required to make proper estimation feasible follows from the analysis from [24] dealing with intrinsic parameter estimation for cameras. Translating the results of that paper to the problem at hand, it follows that observations at two different poses provide sufficient constraints to avoid depth ambiguity, and at each pose, at least four LEDs should be observed to have enough information for the calibration.

In this paper, we estimated the intrinsic parameters from the normalized differences instead of the quad-tuple of RSS values, where the PDF of the normalized differences was approximated by a Gaussian distribution. To evaluate the effect of both the approximation and normalized differencing, we compare the misspecified Cramér-Rao bound $\text{MCRB}(\Theta)$ (30) with the Cramér-Rao bound for the quad-tuple of RSS values $\check{\mathbf{y}}$ observed by the QPD, i.e., $\text{CRB}(\Theta)$, corresponding to the true PDF $p(\check{\mathbf{y}})$. As $\check{\mathbf{y}}$ is Gaussian distributed, the derivation of the Fisher information matrix $\mathbf{J}(\Theta)$ and thus $\text{CRB}(\Theta) = \mathbf{J}^{-1}(\Theta)$ [30] is straightforward and yields

$$\mathbf{J}(\Theta) = (\nabla_{\Theta} \check{\mathbf{g}})^T \check{\Sigma}_{\check{\mathbf{y}}}^{-1} (\nabla_{\Theta} \check{\mathbf{g}}), \quad (36)$$

where $\nabla_{\Theta} \check{\mathbf{g}} = [\nabla_{\Theta} \mathbf{g}_{1,1}, \nabla_{\Theta} \mathbf{g}_{1,2}, \dots, \nabla_{\Theta} \mathbf{g}_{N_T, N_L}]^T \in \mathbb{R}^{(4N_T N_L) \times (6N_T + 3)}$ denotes the gradient of $\check{\mathbf{g}}$ with respect to Θ .

3.3. Pose Estimation

Once the intrinsic parameters are estimated during the calibration process, the estimated parameters can be used for accurate pose estimation of the receiver. To estimate the pose, a similar procedure can be used for the calibration by means of the Gauss–Newton method, except that now it is assumed that the intrinsic parameters \mathbf{k} are known. At pose i , the observation \mathbf{t}_i only depends on the pose \mathbf{T}_i and is independent of other poses. Therefore, we can estimate the poses independently, in contrast to the calibration procedure that needed to consider the observations of multiple poses due to the presence of the unknown intrinsic parameters. Again, the estimation is obtained with an iterative procedure of which the update step is the same as in (28):

$$\mathbf{T}_i^{t+1} = \exp\left(\left(\tau_{\mathbf{T}} \Delta_{\mathbf{T}_i}(\mathbf{k}, \mathbf{T}_i^t)\right)^\wedge\right) \mathbf{T}_i^t \quad (37)$$

Similar to the calibration procedure, we can obtain a coarse estimate by means of the direct linear transformation method together with (29) to initialize the Gauss–Newton method for pose estimation.

To evaluate the pose estimates, the position error is expressed by the Euclidean distance between the position vector \mathbf{r} and its estimate $\hat{\mathbf{r}}$, i.e., $\mathbf{r}_e = \hat{\mathbf{r}} - \mathbf{r}$, while the orientation error is expressed by the axis-angle vector between the orientation matrix \mathbf{R} and its estimate $\hat{\mathbf{R}}$, i.e., $\mathbf{u}_e = (\log(\hat{\mathbf{R}}\mathbf{R}^T))_{\otimes}$ [27]. Similar to the pose estimator derived from the calibration

algorithm, we can derive the theoretical lower bounds for the pose estimate, i.e., $\text{MCRB}(\mathbf{T}_i)$ and $\text{CRB}(\mathbf{T}_i)$, from (35) and (36), respectively, by restricting the parameters Θ to \mathbf{T}_i . Then the left upper and right bottom 3×3 submatrices of $\text{MCRB}(\mathbf{T}_i)$ are bounds on position and orientation, respectively. The same result applies to $\text{CRB}(\mathbf{T}_i)$.

4. Numerical Assessment

In this section, we first verify the performance of the proposed calibration algorithm and compare the mean squared error with the theoretical bound based on simulations. Then, the calibration procedure is used to determine the intrinsic parameters in a simulation setup, after which the pose is estimated. To obtain a comparative study, this section gives a comparison between our proposed pose estimation method and the camera's PnP and iterative methods.

4.1. Calibration

As the number of parameters to be estimated in the calibration phase is large, to have good performance, the number and quality of observations must be sufficiently high. To ensure that these concerns are met, we consider a dedicated calibration setup, where we place the LEDs in a plane at vertical distance sufficiently close to the receiver, i.e., $z = 0.65$ m, to obtain a sufficiently high SNR. Furthermore, to guarantee the receiver has sufficient LEDs within its FOV, we assume $N_L = N_{L,r}^2$ LEDs are placed in a square grid with an area of 400 cm^2 , where the LED in the i th row and j th column has position $\mathbf{r}_L = \left(\frac{20}{N_{L,r}-1} \begin{bmatrix} i-1 \\ j-1 \\ 0 \end{bmatrix} - \begin{bmatrix} 10 \\ 10 \\ 0 \end{bmatrix} \right) \text{ cm}$, $i, j \in \{1, \dots, N_{L,r}\}$. All LEDs are assumed to have a transmitted power $P_t = 1 \text{ W}$, and a Lambertian pattern with order $\gamma = 1$. We set the origin of the system frame as the center of the LED plane.

The receiver consists of a circular QPD with radius $r_p = 5 \text{ mm}$, resulting in an active area $A_p = \pi r_p^2$, and responsivity $R_p = 0.4 \text{ A/W}$. Above the QPD, we place at a height $|h| = 3.0 \text{ mm}$ a circular aperture with radius $l = 2.5 \text{ mm}$. In our simulations, we set the misalignment between the centers of the aperture and the QPD equal to $\mathbf{k}_{m,q} = [0.5, -0.3] \text{ mm}$. The noise in the different quadrants of the QPD is assumed to have a correlation coefficient $\rho = 0.7$. The orientation \mathbf{R} of the receiver is defined with the ZXZ Euler angles $[\theta_\alpha, \theta_\beta, \theta_\gamma]$ [31]. We assume the azimuth and roll angles, i.e., θ_α and θ_γ , are uniformly distributed over $[0, 2\pi)$, and we set the elevation angle θ_β as a variable. In this way, the receiver has a random orientation but a controlled tilt angle of θ_β . After fixing the orientation, we still need to specify the position of the receiver. To ensure that most LEDs are observed, we select the position of the receiver so that it is pointing toward the center of the LEDs plane. This corresponds to the position $\mathbf{r} = \frac{z\mathbf{Re}_3}{e_3^T \mathbf{Re}_3}$.

In the simulations, we use (15) to generate the observations. We assume the shot noise has power spectral density $N_0 = 2.10 \times 10^{-22} \text{ A}^2/\text{Hz}$, which corresponds to a background spectral irradiance $p_n = 5.8 \times 10^{-6} \text{ W}/(\text{cm}^2 \cdot \text{nm})$ and a visible light bandwidth $\Delta\lambda = 360 \text{ nm}$ [32]. Supposing the electrical bandwidth equals $B = 1 \text{ MHz}$ [8], we can compute the noise variance σ_{w3}^2 with $\sigma_{w3}^2 = N_0 B$. Simulation results will be given in terms of the effective SNR $= \left(\frac{(\gamma+1)A_a R_t P_t}{2\pi\sigma_{w3}} \right)^2$ with $A_a = \pi l^2$. (As the received SNR depends on the receiver pose, a fixed received SNR will limit the parameter space of the receiver pose. As this will complicate the simulation, we chose to fix the effective SNR) Given the optical transmitted power $P_t = 1 \text{ W}$, we obtain the effective SNR = 44.73 dB. In our simulations, we will use the range SNR $\in [25, 65] \text{ dB}$ to take into account variations of the system parameters.

First, we evaluate the performance of the designed calibration estimator as a function of SNR. As a baseline method, we consider the calibration algorithm for cameras [33]. Because the algorithm of [33] takes the image of an object as input, which in this case corresponds to the observation of the light spot's position, we use (29) prior to this algorithm to convert the observation \mathbf{t} to the estimated center of the light spot. The number of LEDs in

the LED plane is set to $N_L = 25$, and the receiver observes the plane from $N_T = 4$ randomly generated poses with $\theta_\beta = \pi/9$ rad. We vary the SNR from 25 to 65 dB and plot in Figure 5a the resulting root mean square error (RMSE) (denoted by CAL) versus $\bar{\xi}$, i.e., the received SNR averaged over the four poses. The root of the MCRB and CRB (denoted by rMCRB and rCRB) from Section 3.2 and the performance of the baseline method (denoted by BAS) are also plotted for comparison. As we can see from Figure 5a, at high $\bar{\xi}$, the two calibration methods perform close to optimal, while at low $\bar{\xi}$, the proposed calibration algorithm outperforms the baseline method. This can be explained as the bias in (29) for the baseline method is larger than for the proposed calibration method when $\bar{\xi}$ is low. The gap between MCRB and CRB reflects the net performance cost of the followed strategy, where we use the normalized differences instead of the RSS values and the approximated PDF. The gap shows that the robustness to imperfect knowledge of the transmitted power and radiation pattern of the LEDs, obtained by this strategy, is achieved at the cost of accuracy. The performance loss for the estimation of h is larger than for the estimation of \mathbf{k}_m .

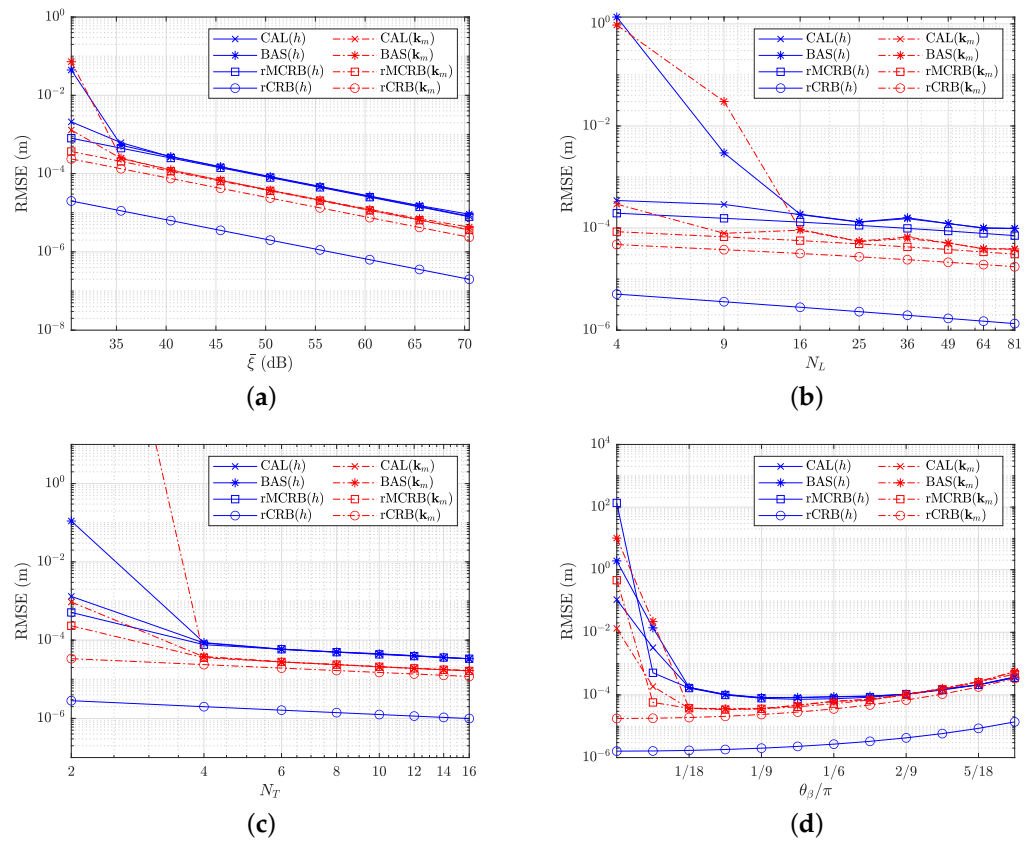


Figure 5. RMSE of the estimator and the theoretical bound: (a) $N_L = 25$, $N_T = 4$, $\theta_\beta = \pi/9$ rad; (b) SNR = 45 dB, $N_T = 3$, $\theta_\beta = \pi/9$ rad; (c) SNR = 45 dB, $N_L = 25$, $\theta_\beta = \pi/9$ rad; (d) SNR = 45 dB, $N_L = 25$, $N_T = 4$.

Next, we evaluate the effect of the number N_L of LEDs and the number N_T of poses. We set SNR = 45 dB and $\theta_\beta = \pi/9$ rad. Both variables influence the number of observations, so we expect that increasing either N_L or N_T will improve the performance of the calibration. In Figure 5b, we show the performance for $N_{L,r} = \sqrt{N_L} = \{2, 3, \dots, 9\}$ with $N_T = 3$, and in Figure 5c for $N_T = \{2, 4, \dots, 16\}$ with $N_L = 25$. The figures show that the lower bound is reached for both the proposed calibration method and the baseline method when N_L or N_T is sufficiently large. In that case, the performance improves logarithmically with N_L and N_T . When N_L or N_T are small, i.e., when the number of observations is small, the proposed method outperforms the baseline method.

Finally, we evaluate the effect of the orientation θ_β of the receiver with respect to the LED plane. We set SNR = 45 dB, $N_L = 25$, and $N_T = 4$. The calibration performance for θ_β varying from 0 rad to $5\pi/18$ rad is shown in Figure 5d. When θ_β is small, the performance first improves by increasing θ_β . This can be explained as for small θ_β , the randomly generated poses are very close to each other implying the solution set is tightly coupled and thus more susceptible to noise. For large θ_β , the performance starts to degrade, as in our assumptions we assumed that the receiver is pointing to the center of the LEDs plane, implying large θ_β corresponds to placing the receiver far from the center of the LEDs plane. As a consequence, the RSS values will experience lower SNR values, causing the slight degradation of the performance. The optimal performance will therefore occur at intermediate values of θ_β .

4.2. Pose Estimation

In the pose estimation phase, we assume the intrinsic parameters are known, i.e., in practice, we use the intrinsic values obtained in the calibration phase. As in this case, we only need to estimate the six parameters of the pose; the number of LEDs can be lower than for the calibration phase. Furthermore, we assume the VLP system is combined with the illumination system, so the number and placement of the LEDs must satisfy the illumination standards. To this end, we consider in the evaluation of the pose estimator a $5\text{ m} \times 5\text{ m} \times 3\text{ m}$ area, where $N_L = 25$ LEDs with optical power $P_t = 5\text{ W}$ are mounted at the ceiling of the area, as shown in Figure 6, unless specified otherwise (assuming the optical power of the LEDs equals 5 W, an average horizontal illuminance of 500 lm (corresponding to an office area) in a plane at $z = 2\text{ m}$ below the LEDs requires roughly 25 LEDs). We define the boundary vector $\mathbf{b} = [5, 5, 3]^T$, the number of LED columns in the X direction $N_{L,X} = 5$ and Y direction $N_{L,Y} = 5$, and the positions of the LEDs are given by $\left[\frac{[\mathbf{b}]_1(2i-1)}{2N_{L,X}}, \frac{[\mathbf{b}]_2(2j-1)}{2N_{L,Y}}, [\mathbf{b}]_3 \right]^T$, with $i \in \{1, \dots, N_{L,X}\}$ and $j \in \{1, \dots, N_{L,Y}\}$, where $[\mathbf{b}]_i$ denotes the i th component of \mathbf{b} . We assume the algorithm only uses the LEDs that result in all quadrants to produce a sufficiently large RSS output to ensure that the information in the RSS values is reliable. To take into account the effect of calibration and the possible uncertainty in the estimated parameters, we assume that the receiver has been calibrated and that the estimated intrinsic parameter vector $\hat{\mathbf{k}} \sim \mathcal{N}(\mathbf{k}, \Sigma_{\hat{\mathbf{k}}})$, with $\Sigma_{\hat{\mathbf{k}}} = \text{diag}([\sigma_{uv}^2, \sigma_{uv}^2, \sigma_h^2])$, and σ_{uv} and σ_h are set to the RMSE of the calibration given in the Section 4.1.

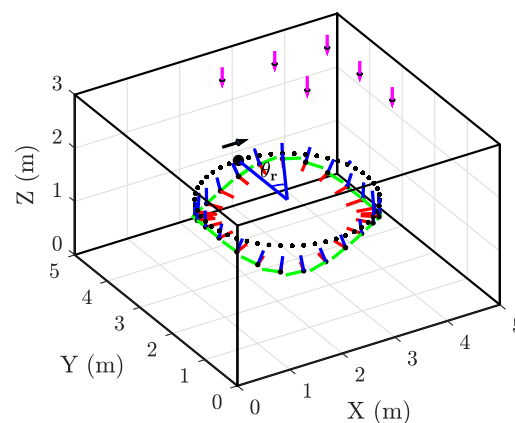


Figure 6. Simulation setup for the QADA receiver. The three orthonormal vectors in three different colors (the red, green, and blue vector represent the x -axis, y -axis, and z -axis, respectively) at each sample on the path represent the frame of the receiver. The pink arrows represent the LEDs (only a fraction of them are shown) on the ceiling. θ_r indicates the traveled angle along the dotted circle in the XY plane.

To evaluate the performance of the proposed estimators, we consider a path with a circular pattern in the XY plane and a sinusoidal pattern in the Z direction as shown in

Figure 6. The radius of the circle in the XY plane is 1.5 m, and the amplitude of the sinus is 0.2 m. The circle is centered at $[2.5, 2.5, 1.2]^T$ m. Starting at the coordinates $[4.0, 2.5, 1.2]^T$ m, the path oscillates sinusoidally in the Z direction and completes the path in three periods.

In Figure 7, we evaluate the proposed estimator with respect to the transmitted power. In this simulation, we set $\Sigma_{\hat{\mathbf{k}}}$ equal to the corresponding RMSE values from the calibration when SNR = 55 dB, i.e., $\sigma_{uv} = 1.17 \times 10^{-5}$ m and $\sigma_h = 2.60 \times 10^{-5}$ m. The RMSE of the pose estimate, averaged over the path, is first evaluated in terms of the transmitted power $P_t = (1 + 2m) \times 10^n$ W, where the variables $m \in \{0, \dots, 4\}$ and $n \in \{0, 1\}$ specify the selected power values in the figure. Figure 7 shows this averaged RMSE of the designed estimator based on the least squares method (LSM) (denoted by LSM) and the root of the MCRB (denoted by rMCRB) as a function of the mean received SNR $\bar{\xi}$. In comparison, we also add the performance of the proposed estimator when perfect knowledge of the intrinsic parameters is given (denoted by LSM-K) and when the intrinsic parameters are not calibrated and assumed to be the design values $|h| = 2.8$ mm and $\mathbf{k}_{m,q} = \mathbf{0}$ mm (denoted by LSM-X), while the true values are randomly selected from a Gaussian distribution with as averages the design values and variances given above. The averaged RMSE of the baseline method [33] (denoted by BAS) and the root of the CRB (denoted by rCRB) are also plotted for comparison. As a second baseline method, we also consider the classical PnP method (denoted by PnP). As expected, the performance of the estimator with perfect knowledge of the intrinsic parameters reaches the theoretical lower bound MRCB and improves with the SNR. On the other hand, the performance of the estimators without perfect knowledge of the intrinsic parameters, i.e., calibrated or not, shows an error floor when the received SNR increases. This error floor is caused by the uncertainty in the estimated intrinsic parameters, which becomes dominant when the received SNR is sufficiently large. Obviously, the error floor is much higher in the absence of calibration. Our algorithm strongly outperforms classical PnP, as expected, while its performance is slightly better than the baseline method BAS. The gap between MCRB and CRB implies that the robustness to imperfect knowledge of the transmitted power and the radiation pattern of the LEDs, obtained by the strategy of normalizing differences, is achieved at the cost of accuracy.

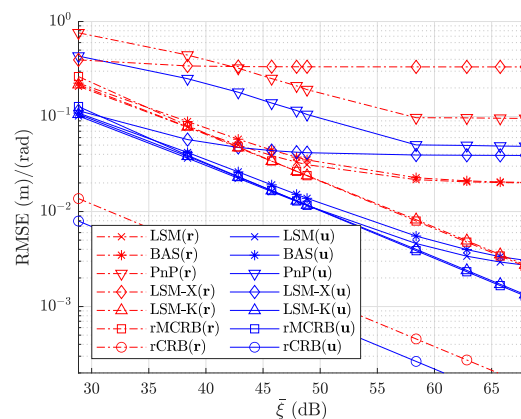


Figure 7. RMSE for position and orientation estimates, compared with the mean received SNR $\bar{\xi}$.

To illustrate the robustness of the proposed algorithm to fluctuations in the transmitted power, we consider the situation where the transmitted power of an LED follows a Gaussian distribution [34], i.e., $P_t = \kappa P_{t,m}$, with $\kappa \sim \mathcal{N}(1, \sigma_\kappa^2)$. Experiments in [34] show that the standard deviation σ_κ in general is smaller than $\sigma_\kappa < 0.0667$. In Figure 8, we plot the RMSE averaged over the entire path as a function of the standard deviation $\sigma_\kappa \in [0, 0.0667]$, assuming $P_{t,m} = 5$ W, $\sigma_{uv} = 1.17 \times 10^{-5}$ m, and $\sigma_h = 2.60 \times 10^{-5}$ m. We observe that for the given range of σ_κ , the RMSE is (essentially) independent of σ_κ , implying the approach of the normalized differences is robust to fluctuations in the transmitted power.

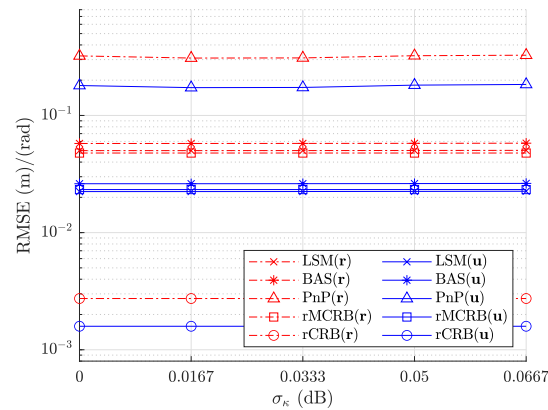


Figure 8. RMSE for position and orientation estimates as a function of σ_{κ} .

Next, we show the RMSE of the pose estimates of the proposed estimator as a function of the uncertainty in $\hat{\mathbf{k}}$ for $P_t = 5$ W and compare the resulting RMSE with that of the baseline method and with the theoretical lower bounds. Figure 9 shows that, as expected, the performance of the estimator with the perfect knowledge of the intrinsic parameters reaches the theoretical lower bound and is independent of the uncertainty. The figure also illustrates that a large uncertainty degrades the performance of the estimators significantly. When the uncertainty becomes smaller, the performance of the estimators of LSM and BAS gradually reach and are bounded by the theoretical lower bound. This is expected, as due to the uncertainty, the estimators are in the best case as good as the ones with perfect knowledge of the intrinsic parameters. (Close to) optimal performance is achieved as long as the uncertainty is sufficiently small.

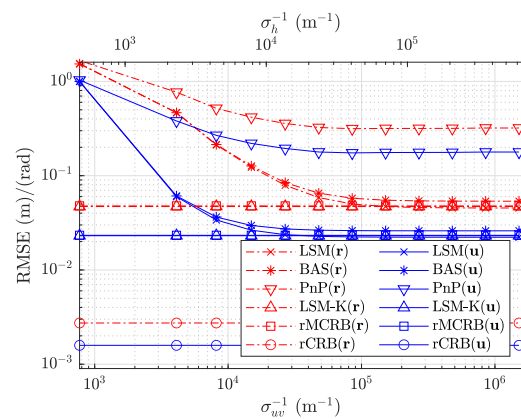


Figure 9. RMSE for position and orientation estimates as a function of σ_{uv} and σ_h .

Finally, we evaluate the performance of the proposed receiver over the whole considered area. For the LEDs, we consider the setup of Figure 6, where the $N_L = 25$ LEDs are attached in a square grid to the ceiling of the $5\text{ m} \times 5\text{ m}$ area. We assume the receiver is placed at the positions belonging to a horizontal grid with a horizontal spacing of 0.2 m, at a vertical distance $z_d \in \{1, 1.5, 2, 2.5, 3\}$ m from the LEDs. Taking into account that the receiver in practice is pointing roughly upwards, we consider for each sample point a random orientation with an elevation angle θ_{β} that is uniformly distributed within the interval $[0, \theta_{\beta,m}]$ rad. We consider two scenarios, i.e., $\theta_{\beta,m} = 0$ rad for the case where the receiver is always pointing upwards, and $\theta_{\beta,m} = \pi/3$ rad for the tilted scenario. Figure 10 shows the cumulative distribution (CDF) of the pose error when $z_d = 2$ m. We observe that the proposed algorithm outperforms the baseline methods BAS and PnP. The figure further reveals that our algorithm converges approximately for 95% of the sample positions when $\theta_{\beta,m} = 0$ rad, i.e., when the receiver points upwards, while this reduces to 75%

when a random tilted angle is considered, i.e., when $\theta_{\beta,m} = \pi/3$ rad. Hence, as expected, the coverage degrades if the receiver has a random tilt. To evaluate the effect of the vertical distance z_d on the performance, we compute the probability that the error is smaller than 1 m for the position and 0.2 rad for the orientation. Figure 11 shows that the performance of our estimator improves when the distance increases, as well as for the baseline method BAS. Despite the larger distance, and thus lower SNR, increasing the distance will bring more LEDs into the field-of-view of the receiver, implying more information is available to estimate the pose. On the other hand, for the PnP method, the performance degrades when the distance increases. This is explained as the PnP method is more sensitive to a reduction of the SNR, resulting in a performance degradation despite the larger number of LEDs available within the field-of-view.

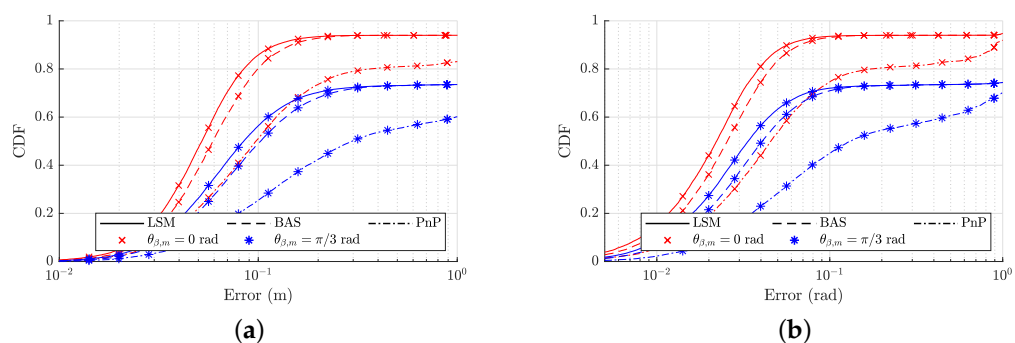


Figure 10. CDF of RMSE for position and orientation estimates for $z_d = 2$ m: (a) position error; (b) orientation error.

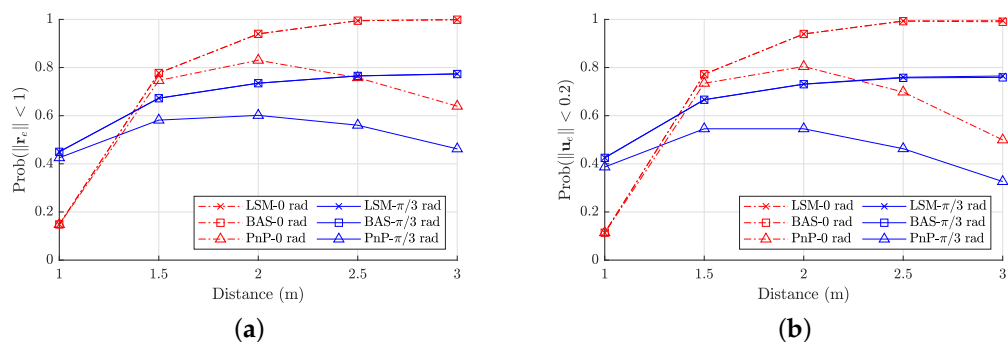


Figure 11. Probability versus the distance z_d for: (a) position error $\|\mathbf{r}_e\| < 1$ m; (b) orientation error $\|\mathbf{u}_e\| < 0.2$ rad.

5. Conclusions

In this paper, we investigate the pose estimation and calibration problem for the QADA receiver. To this end, the channel link is first modeled in terms of the receiver's pose and misalignment using the perspective projection model. Then, we study the effect of the strategy of differencing and normalization to the probability density function (PDF), and it is shown that the applied strategy results in a complex PDF. Therefore, we derive an approximated PDF in a simple closed-form, based on which the calibration and the pose estimation algorithms using the least squares principle are proposed. The proposed calibration does not require any information about the pose of the receiver and is robust to the received power variation and imperfect knowledge of the LEDs' radiation patterns, while the proposed pose estimator is able to directly estimate the pose from the (normalized) RSS values. Computer simulations confirm that the proposed calibration algorithm is effective and that the proposed pose estimator outperforms the PnP algorithm.

Author Contributions: S.S. led the paper, drafted the main parts, derived the estimation methods; J.M.M.S. has supported the work on verification of the estimation methods; S.L. and H.S. were supervisors of the research and acquired the funding; H.S. proofread and edited the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially funded by the EOS grant 30452698 from the Belgian Research Councils FWO and FNRS, by the Flemish Government (AI Research Program), by the National Natural Science Foundation of China under Grant 61771474, by the National Secretariat of Higher Education, Science, Technology and Innovation of Ecuador (SENESCYT).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The study did not report any data.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. The Expressions for the Gradient $\nabla_{\Theta}\mu_{i,j}$

Based on the chain rule, $\nabla_{\Theta}\mu_{i,j} = \nabla_{\mathbf{r}_{S,q,i,j}}\mu_{i,j}\nabla_{\Theta}\mathbf{r}_{S,q,i,j} \in \mathbb{R}^{2 \times (6N_T+3)}$, where $\nabla_{\mathbf{r}_{S,q,i,j}}\mu_{i,j}$ is the gradient of $\mu_{i,j}$ with respect to $\mathbf{r}_{S,q,i,j}$, and $\nabla_{\Theta}\mathbf{r}_{S,q,i,j}$ is the gradient of $\mathbf{r}_{S,q,i,j}$ with respect to Θ . $\nabla_{\mathbf{r}_{S,q,i,j}}\mu_{i,j}$ is determined by its k^{th} component

$$\left[\nabla_{\mathbf{r}_{S,q,i,j}}\mu_{i,j}\right]_k = \frac{(2 - 2\cos 2\alpha_k)\mathbf{e}_k^T}{l\pi\sqrt{1 - (\mathbf{e}_k^T\mathbf{r}_{S,q,i,j}/l)^2}}, \quad (\text{A1})$$

while $\nabla_{\Theta}\mathbf{r}_{S,q,i,j} = \left[\nabla_{\mathbf{K}}\mathbf{r}_{S,q,i,j}, \nabla_{\mathbf{T}_1}\mathbf{r}_{S,q,i,j}, \dots, \nabla_{\mathbf{T}_{N_T}}\mathbf{r}_{S,q,i,j}\right] \in \mathbb{R}^{2 \times (6N_T+3)}$ is given by

$$\nabla_{\mathbf{K}}\mathbf{r}_{S,q,i,j} = \frac{(\mathbf{T}_i\bar{\mathbf{r}}_{L,j})^{\odot}}{\mathbf{e}_3^T\mathbf{T}_i\bar{\mathbf{r}}_{L,j}} \quad (\text{A2})$$

and

$$\nabla_{\mathbf{T}_k}\mathbf{r}_{S,q,i,j} = \begin{cases} \left(\frac{\mathbf{K}}{\mathbf{e}_3^T\mathbf{T}_i\bar{\mathbf{r}}_{L,j}} - \frac{\mathbf{K}\mathbf{T}_i\bar{\mathbf{r}}_{L,j}\mathbf{e}_3^T}{(\mathbf{e}_3^T\mathbf{T}_i\bar{\mathbf{r}}_{L,j})^2}\right)(\mathbf{T}_i\bar{\mathbf{r}}_{L,j})^{\odot} & \text{if } k = i \\ \mathbf{0} \in \mathbb{R}^{2 \times 6} & \text{if } k \neq i \end{cases}, \quad (\text{A3})$$

where the gradient of $\mathbf{r}_{S,q,i,j}$ with respect to \mathbf{T}_k (A3) is calculated with the infinitesimal perturbation, and the operators \cdot^{\odot} and \cdot^{\ominus} are defined as

$$\left([a_1, a_2, a_3, a_4]^T\right)^{\odot} = \begin{bmatrix} a_3 & 0 & a_1 \\ 0 & a_3 & a_2 \end{bmatrix}, \quad (\text{A4})$$

and

$$\left([\xi^T, \eta]^T\right)^{\ominus} = \begin{bmatrix} \eta\mathbf{I}_3 & -\xi \times \\ \mathbf{0}^T & \mathbf{0}^T \end{bmatrix}, \quad (\text{A5})$$

respectively.

References

- Hassan, N.U.; Naeem, A.; Pasha, M.A.; Jadoon, T.; Yuen, C. Indoor Positioning Using Visible LED Lights: A Survey. *ACM Comput. Surv.* **2015**, *48*, 1–32. [[CrossRef](#)]
- Zhang, X.; Gao, Q.; Gong, C.; Xu, Z. User grouping and power allocation for NOMA visible light communication multi-cell networks. *IEEE Commun. Lett.* **2016**, *21*, 777–780. [[CrossRef](#)]
- Alsaadi, F.E.; Elmirghani, J.M. High-speed spot diffusing mobile optical wireless system employing beam angle and power adaptation and imaging receivers. *J. Light. Technol.* **2010**, *28*, 2191–2206. [[CrossRef](#)]
- Eroglu, Y.S.; Anjinappa, C.K.; Guvenc, I.; Pala, N. Slow Beam Steering and NOMA for Indoor Multi-User Visible Light Communications. *IEEE Trans. Mob. Comput.* **2021**, *20*, 1627–1641. [[CrossRef](#)]

5. Eroglu, Y.S.; Guvenc, I.; Sahin, A.; Pala, N.; Yuksel, M. Diversity combining and piezoelectric beam steering for multi-element VLC networks. In Proceedings of the 3rd Workshop on Visible Light Communication Systems, New York, NY, USA, 3–7 October 2016; pp. 25–30.
6. Rahaim, M.B.; Morrison, J.; Little, T.D.C. Beam Control for Indoor FSO and Dynamic Dual-Use VLC Lighting Systems. *J. Commun. Inf. Netw.* **2017**, *2*, 11–27. [[CrossRef](#)]
7. Beysens, J.; Wang, Q.; Van Den Abeele, M.; Pollin, S. BlendVLC: A Cell-free VLC Network Architecture Empowered by Beamspot Blending. In Proceedings of the IEEE Conference on Computer Communications, INFOCOM 2021, Vancouver, BC, Canada, 10–13 May 2021; pp. 1–10.
8. Cincotta, S.; He, C.; Neild, A.; Armstrong, J. High angular resolution visible light positioning using a quadrant photodiode angular diversity aperture receiver (QADA). *Opt. Express* **2018**, *26*, 9230–9242. [[CrossRef](#)]
9. Cincotta, S.; He, C.; Neild, A.; Armstrong, J. Indoor Visible Light Positioning: Overcoming the Practical Limitations of the Quadrant Angular Diversity Aperture Receiver (QADA) by Using the Two-Stage QADA-Plus Receiver. *Sensors* **2019**, *19*, 956. [[CrossRef](#)]
10. Aparicio-Esteve, E.; Hernández, A.; Ureña, J.; Villadangos, J.M. Visible Light Positioning System based on a Quadrant Photodiode and Encoding Techniques. *IEEE Trans. Instrum. Meas.* **2019**, *69*, 5589–5603. [[CrossRef](#)]
11. Aparicio-Esteve, E.; Hernández, A.; Ureña, J.; Villadangos, J.M.; Ciudad, F. Estimation of the Polar Angle in a 3D Infrared Indoor Positioning System based on a QADA receiver. In Proceedings of the 2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Pisa, Italy, 30 September–3 October 2019; pp. 1–8.
12. Aparicio-Esteve, E.; Hernández, A.; Ureña, J. Design, Calibration, and Evaluation of a Long-Range 3-D Infrared Positioning System Based on Encoding Techniques. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–13. [[CrossRef](#)]
13. Mohammed, M.; He, C.; Cincotta, S.; Neild, A.; Armstrong, J. Communication aspects of visible light positioning (VLP) systems using a quadrature angular diversity aperture (QADA) receiver. *Sensors* **2020**, *20*, 1977. [[CrossRef](#)]
14. Aparicio-Esteve, E.; Ureña, J.; Hernández, A.; Pizarro, D.; Moltó, D. Using Perspective-n-Point Algorithms for a Local Positioning System Based on LEDs and a QADA Receiver. *Sensors* **2021**, *21*, 6537. [[CrossRef](#)] [[PubMed](#)]
15. Lepetit, V.; Fua, P. *Monocular Model-Based 3D Tracking of Rigid Objects*; Now Publishers: Norwell, MA, USA, 2005.
16. Lin, B.; Tang, X.; Ghassemlooy, Z.; Li, Y.; Zhang, S.; Wu, Y.; Li, H. An indoor VLC positioning system based on OFDMA. In Proceedings of the Asia Communications and Photonics Conference 2016, Wuhan, China, 2–5 November 2016.
17. Lin, B.; Tang, X.; Ghassemlooy, Z.; Lin, C.; Li, Y. Experimental demonstration of an indoor VLC positioning system based on OFDMA. *IEEE Photonics J.* **2017**, *9*, 1–9. [[CrossRef](#)]
18. Steendam, H.; Wang, T.Q.; Armstrong, J. Theoretical Lower Bound for Indoor Visible Light Positioning Using Received Signal Strength Measurements and an Aperture-Based Receiver. *J. Light. Technol.* **2017**, *35*, 309–319. [[CrossRef](#)]
19. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*; Cambridge University Press: Cambridge, UK, 2003.
20. Kahn, J.M.; Barry, J.R. Wireless infrared communications. *Proc. IEEE* **1997**, *85*, 265–298. [[CrossRef](#)]
21. Marsaglia, G. Ratios of normal variables. *J. Stat. Softw.* **2006**, *16*, 1–10. [[CrossRef](#)]
22. Curtiss, J.H. On the Distribution of the Quotient of Two Chance Variables. *Ann. Math. Stat.* **1941**, *12*, 409–421. [[CrossRef](#)]
23. Hayya, J.; Armstrong, D.; Gressis, N. A note on the ratio of two normally distributed variables. *Manag. Sci.* **1975**, *21*, 1338–1341. [[CrossRef](#)]
24. Zhang, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334. [[CrossRef](#)]
25. Nguyen, H.T.; Kreinovich, V. *Applications of Continuous Mathematics to Computer Science*; Springer: Berlin/Heidelberg, Germany, 1997; Volume 38.
26. Absil, P.A.; Mahony, R.; Sepulchre, R. *Optimization Algorithms on Matrix Manifolds*; Princeton University Press: Princeton, NJ, USA, 2009.
27. Shen, S.; Li, S.; Steendam, H. Simultaneous Position and Orientation Estimation for Visible Light Systems with Multiple LEDs and Multiple PDs. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 1866–1879. [[CrossRef](#)]
28. Fortunati, S.; Gini, F.; Greco, M.S.; Richmond, C.D. Performance Bounds for Parameter Estimation under Misspecified Models: Fundamental Findings and Applications. *IEEE Signal Process. Mag.* **2017**, *34*, 142–157. [[CrossRef](#)]
29. Rasmussen, C.E. Gaussian processes in machine learning. In *Summer School on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 63–71.
30. Kay, S.M. *Fundamentals of Statistical Signal Processing: Estimation Theory*; Prentice-Hall: Upper Saddle River, NJ, USA, 1993.
31. Pissanetzky, S. *Rigid Body Kinematics and C++ Code*; SciControls.com: Houston, TX, USA, 2005.
32. Steendam, H. A 3-D Positioning Algorithm for AOA-Based VLP With an Aperture-Based Receiver. *IEEE J. Sel. Areas Commun.* **2018**, *36*, 23–33. [[CrossRef](#)]
33. Bouguet, J.Y. Camera Calibration Toolbox for Matlab. 2008. Available online: http://www.vision.caltech.edu/bouguetj/calib_doc (accessed on 23 June 2022).
34. Plets, D.; Bastiaens, S.; Martens, L.; Joseph, W.; Stevens, N. On the impact of LED power uncertainty on the accuracy of 2D and 3D Visible Light Positioning. *Optik* **2019**, *195*, 163027. [[CrossRef](#)]