

Article

# A Novel Ranging Technique Based on Optical Camera Communications and Time Difference of Arrival

Jose Rabadan <sup>\*,†,‡</sup> , Victor Guerra <sup>‡</sup> , Carlos Guerra <sup>‡</sup>  and Julio Rufo <sup>‡</sup>  
and Rafael Perez-Jimenez <sup>‡</sup> 

Institute for Technological Development and Innovation in Communications (IDeTIC),  
Universidad de Las Palmas de Gran Canaria, 35017 Las Palmas de Gran Canaria, Spain;  
vguerra@idetec.eu (V.G.); carlos.guerra@fpct.ulpgc.es (C.G.); jrufo@idetec.eu (J.R.); rafael.perez@ulpgc.es (R.P.-J.)

\* Correspondence: jrabadan@idetec.eu; Tel.: +34-928-45-99-14

† Current address: Edificio Polivalente II, Campus de Tafira, Las Palmas de Gran Canaria,  
35017 Las Palmas, Spain.

‡ These authors contributed equally to this work.

Received: 3 May 2019; Accepted: 5 June 2019; Published: 11 June 2019



**Abstract:** In this work, a new Time Difference of Arrival (TDoA) scheme for distance measurement based on Optical Camera Communication (OCC) systems is proposed. It relies on the use of optical pulses instead of radio-frequency signals as the time reference triggers, and the introduction of a rolling shutter camera, whose characteristics allows substituting the timer modules used in conventional TDoA techniques by image processing of the illuminated area in the picture. This processing on the camera's images provides time measurements and implies a specific analysis, which is presented in this work. The system performance and properties, such as resolution and range, mainly depends on the camera characteristics, such as the frames capture rate and the image quality. This new technique is suitable to be implemented in smartphones or other Commercial Off-The-Shelf (COTS) devices equipped with a camera and speakers.

**Keywords:** Optical Camera Communications; ranging technique; statistical pulse detection

## 1. Introduction

Positioning systems are gaining an essential role in indoor scenarios, with several applications on accessibility and marketing [1,2]. In addition to classic Visible Light Positioning (VLP) indoor scenarios (commercial areas, hospitals, and visually impaired-people guidance [3]), applications of this technology can be also oriented to location on industrial environments, (e.g., robot positioning) where the presence of high electromagnetic (EM) noise levels induce huge errors when using traditional Radio Frequency (RF) systems [4]. Another scenario is underground mining, especially in the case of potentially flammable gases emission, as in coal extraction facilities (or even when coal is just present) or when there is a risk of the atmosphere becomes potentially explosive (e.g., oil and gas production or storage plants). Optical Camera Communication (OCC) and LED-based positioning is an alternative to avoid EM emissions [5].

Most of the techniques proposed until present are based on radio frequency signals, generated specifically for that purpose [6], or reused from other systems such as WiFi network devices [7]. Nevertheless, there are also some techniques based on optical signals, which make use of the Solid State Illumination (SSL) devices present in LED lamps [8,9]. This alternative introduces a double purpose of lights in facilities: as light sources and as part of indoor location systems.

Positioning techniques mainly consist of distance measurements among the target to be located, and several reference points to carry out a trilateration operation [1]. Several schemes for distance

estimation have been proposed such as Received Signal Strength (RSS), Time of Flight (ToF), Angle of Arrival (AoA) or Time Difference of Arrival (TDoA) [6]. In this work, a TDoA scheme based on a Visible Light Communication (VLC) measurement system, as previously proposed in [10], is presented. The main contribution of this work is the use of OCC techniques and image processing for measuring the signal's time of flight between the base station (BS) and the mobile station (MS).

Some authors have proposed Optical Camera Communication for Visible Light Positioning (VLP). They are based on the extraction of different information from the camera's images and, in general, use them for estimating distances among reference points and mobile nodes [11,12]. Additionally, the use of OCC to carry out node identification (reference and/or target) has also been studied [13–16], achieving centimeter accuracy in small-room scenarios. Thus, the system can obtain both position and identification from the nodes. In this work, OCC technology is involved in the ranging process rather than in the trilateration one, and makes use of the images in a very different way, which is shown in this paper.

The proposed method reduces the time resolution of the system, since it is limited by the camera scanning rate, but improves the received optical signal detection due to the camera's sensitivity. Furthermore, this scheme can be easily implemented over smartphones using the flash LED (optical emitter), the speaker (sound emitter), and the camera (OCC receiver). The remaining elements are allocated using an external module (optical and acoustic receivers, and optical emitter) unless they can also be integrated into another smartphone: embedded camera as the optical receiver, microphone as the sound receiver, and flash LED as the optical emitter.

The presented work is focused on a new distance estimation technique and describes in depth the proposed distance-measurement scheme. The solutions implemented for adapting the signal transmission and processing, as well as the error performance are also discussed. Several of the presented modules are required for implementing a full positioning system, which are integrated following the general trilateration scheme used by other indoor location alternatives. This work also explores a novel error-mitigation strategy based on the shape of the probability density function associated with the ranging procedure.

This paper is organized as follows. First, a thorough analysis of the current context in VLP is provided in Section 2. In Section 3, a description of the proposed method for obtaining the distance is presented. Section 4 analyzes the proposed method concerning performance. Section 5 discusses some relevant statistical aspects of the measurement technique. In Section 6, some relevant results are presented. Finally, Section 7 provides a thorough discussion of the results and their potential impact on the field.

## 2. Brief Literature Review

This section tries to contextualize the presented work introducing the state-of-the-art VLP solutions. The main contributions in the field have been organized with respect to the type of receiver: photodiode or camera. Photodiode-based VLP has been extensively studied, especially at the beginning of Visible Light Communication (VLC). However, camera-based approaches, which use OCC for both ID detection and position estimation are currently being profoundly studied.

### 2.1. Photodiode-Based Visible Light Positioning

Most contributions found in the literature regarding VLP address applications of the technology and the use of different techniques.

Dong et al. reported a fully integrated transmitter using 180 nm BIPOLAR-CMOS-DMOS (BCD) integrated circuits technology [17]. For VLC and VLP applications to get into the mass market, integrated circuit based System on Chip (SoC) and System in a Package (SiP) is the only solution in the real world. The proposed transmitter is based on a Manchester-encoded 12 MHz signal. Although the authors claimed the feasibility of their implementation for VLP, they offered no application results.

An Indoor Positioning System (IPS) using LED identification technology was proposed [18] by Lou et al. LED sources transmit unique ID codes using 200 kbps OOK and the positioning is cell-based, achieving poor accuracies around 1.3 m.

Do et al. evaluated a TDoA-based technique applied to VLP [19]. They demonstrated that it is possible to perform the position estimation without the need for lamp identification but using a pure pilot signal instead. Nevertheless, the pilot signal time requirements (10 ns pulses in a periodic 50 ns signal) makes this scheme unfeasible using conventional lighting LED lamps.

VLP has also been tested on vehicles. An example was presented by Roberts et al. [20], who proposed a phase-difference-of-arrival scheme for estimating the distance between the leading and following vehicles. The system employs the leading car tail lights as emitters of amplitude modulated ranging tones, and photodiodes in the next car headlight for detecting those tones. Although the authors presented their work as VLP, this system is not a positioning system but a ranging scheme to obtain the relative position between vehicles.

An example of an actual VLP system for vehicles was presented [21] by Bai et al. They proposed a positioning technique based on TDoA measurements of signals transmitted by traffic lights. Methods for VLP with one traffic light and two traffic lights are proposed. This work and the previous one present the problem of detecting differences in the received optical signal by two photodiodes separated by a short distance (the vehicle's headlights distance), because of the high propagation speed of light. This problem requires the use of more complex and faster electronic devices, as well as synchronized sampling.

Chen et al. described the demodulation characteristics of an optical receiver using a microphone jack [22]. They found that the communication is feasible and that the smartphone can decode the lamp ID. An Indoor Positioning System (IPS) based on this VLC system is proposed and demonstrated. Nevertheless, no accuracy data are provided, but it could be tens of centimeters since this method only provides relative positioning based on the beacon data frame that the LED lamp send.

Chen and Lain explored VLP with z-axis restriction to provide xy-coordinates [23]. The system retrieves both ID and relative emitter coordinates using OCC under on off keying (OOK), and then applied closed-form equations to provide the xy estimate. The experiments were carried out at 1 m height with an accuracy of around 3 cm. Generally, IPS based on visible light assume that receivers are pointing upwards. Some authors have explored the inclusion of the receiver's attitude estimation into the positioning system with relative success.

The estimated position may present a significant error when the receiver is tilted. In the work presented by Jeong et al. [24], a positioning estimation compensation technique with known tilting angle is presented. Feasibility is verified both theoretically and experimentally, achieving accuracies around 1.5 cm at 100 cm.

In [25], the authors evaluated the performance of a multi-photodiode receiver in a Received Signal Strength Indicator (RSSI) positioning system. They used the relative position of each photodiode in the receiver and the coordinates of the reference light sources. A centimeter accuracy is reported for realistic scenarios.

RSSI-based positioning was theoretically studied by Zhang et al. in [26], concluding that its performance depends on the layout of the LED lamps, AoA, type of light source, and receiver characteristics.

In [27], Li et al. proposed a fusion positioning system based on extended Kalman filters, which fuse the VLC position and the inertial navigation data. The VLC-based positioning data are acquired in an LED lighting environment with a trilateral RSS-based algorithm. The accuracy of the combined positioning system is better than the VLC-based positioning or inertial navigation alone, as expected. The fusion algorithm provides an improvement around 57% respect to VLC positioning, achieving centimeter accuracy.

Some authors have addressed the problem of medium access both theoretically and experimentally. A typical solution is to assign a different Radio Frequency (RF) subcarrier to each anchor node (Frequency Division Multiple Access (FDMA)), but other access techniques have been proposed

(Time Division Multiple Access (TDMA), Orthogonal Frequency Division Multiple Access (OFDMA), Code Division Multiple Access (CDMA), etc.).

Kim et al. implemented a Real Time IPS using RF carrier allocation in VLC [28]. In this work, the authors implemented and tested an RSS-based system in which each LED lamp sends different RF carriers, ranging from 0.1 to 2.1 kHz. The relative positioning error is approximately 7%. The same authors analyzed through simulation the impact of moving speed and receiver's orientation in OOK-based VLP [29], comparing the received optical power with pre-calculated Line of Sight (LOS) channel gains. As could be expected, moving speed and tilting dramatically affect performance.

In [30], the authors analyzed the effect of different multiplexing methods in the performance of VLP systems. FDMA scheme provides the best results in positioning but lowers luminous efficiency. On the other hand, random access procedure shows similarly good behavior but also excellent luminous efficiency. However, the identification sequences transmitted by lights must be short relative to the average delay between two transmissions in order to ensure low collision probability.

A two-phase positioning algorithm was proposed by Prince and Little [31]. During the coarse phase, a positional estimate is returned even if only one anchor node is within the receiver's Field Of View (FOV). The fine phase requires several emitters and its accuracy increases with the number of viewed luminaires. The authors evaluated through simulation the positioning error using different modulations/encodings, medium access strategies and parameter estimation technique (AoA, RSS, ToF). The results suggest that Time Division Multiplexing in combination with Time of Flight measurements provides the best error performance (8 cm).

Another interesting but sophisticated approach to IPS is fingerprinting. In this technique, offline captured data are associated with each position and nonlinear regression machines are trained to fit the data. This approach needs calibration but eases the positioning algorithm from the mobile node's point of view.

In [32], Alam et al. evaluated a Visible Light Positioning system based on RSS. They generated an offline fingerprint database in a  $3.3 \times 2.1$  m floor space illuminated by four luminaires, each one emitting a different square-wave frequency. The fingerprints comprise the RSS of each luminaire, and a weighed K-nearest-neighbors regression is carried out with interesting results. Different metrics are used and the best-achieved accuracy is 2.2 cm. Despite the promising results, the proposed methodology requires prior calibration.

Konings et al. carried out data fusion from RF and VLP systems [33]. In the proposed scheme, the authors used a photodiode-based approach for the VLP subsystem, consisting of four different frequencies ranging from 2 kHz to 4 kHz. The VLP subsystem is very similar to the one presented several years ago by Kim et al. in [28]. The RF subsystem comprises ZigBee-based location hardware. The proposed system requires an offline calibration stage, in which a Convolutional Neural Network is trained using both VLC and RF data. The system exhibits almost perfect accuracy in cell-based location and xy positioning errors around 10 cm for the two tested indoor scenarios. Although the results are interesting, the VLC link samples used during the training phase are obtained with the receiver pointing upwards, which limits the actual usability of the system. Unlike this intrinsic limitation of RSS-based VLP systems, the scheme proposed in this paper is partially immune to misalignment errors.

Hosseiniyanfar et al. [34] based their work on using the uplink channel impulse response to locate the user. This technique takes advantage of the diffuse components of the uplink channel impulse response for positioning. The authors proposed exploiting the line of sight (LOS) component, the second power peak (SPP) of the impulse response, and the delay time between LOS and SPP for positioning. A proof of concept analysis using fixed reference points with uplink photodetectors is presented. Simulation results show a root mean square (RMS) positioning accuracy of 25 cm and 5 cm for one- and four-detector scenarios, respectively.

The effects of multipath propagation in a VLP system based on RSS were studied by Gu et al. [35]. Comparisons with other works where no reflections are considered demonstrate that this phenomenon considerably reduces the positioning accuracy. The authors proposed different calibration approaches

to decrease the effect of multipath reflections, such as the strongest signals selection process or the use of dense lamp distributions.

Huang et al. studied by simulation the use of very low duty cycle Pulse Positioning Modulation (PPM) in VLP [36]. They aimed to keep VLP and VLC functionalities while maintaining a very low illumination level. They ensured the feasibility of the system, but concluded that higher bandwidth would be needed to process the short light impulse signals.

Chen and You described a cooperative system for VLP and VLC based on the spread-spectrum plugin of optical identification (OPID) [37]. By keeping the existing IEEE 802.15.7 [38] frame structure untouched, two different implementation schemes were designed by adding the location information field into the communication information field for VLP applications. It is deemed to be an excellent solution to overcome the mutual exclusion between VLP and VLC systems. No implementation or information about expected accuracy is provided.

Alsalamy proposed the use of MiniMax filters based on Game Theory instead of Kalman filters in VLP [39], based on the better performance of MiniMax when the statistics of noise sources are not available. Results show a better performance of MiniMax when there are disturbances and similar performance of both types when the noise statistics are known.

## 2.2. Camera-Based VLP

As happens in photodiode-based VLP, most contributions are focused on demonstrating the technology. In this case, the systems are based on the transmission of the lamp's Identification (ID) using OCC, and then estimating the position using the geometrical relationships between the captured image and the scenario's geometry.

In [40], Yoshino et al. proposed a VLP system based on image sensors. The system is based on receiving the absolute location information from three LED sources using a camera and estimating the coordinates iteratively as an optimization process. The authors did not comment on how the transmission is carried out, and they reported accuracies in the range of tens of centimeters.

Nakazawa et al. presented [41] a VLP system with LED tracking. The authors used a fisheye-lens-equipped camera with 160° FOV. The LED tracking feature was achieved using optical flow algorithms and the lamp ID transmission was based on a 4-PPM encoding. The authors obtained accuracies around 10 cm in a 100 cm-height scenario. The authors did not comment on whether they used a lens distortion compensation algorithm. Fisheye lenses introduce significant errors at relatively small deviations from the optical axis and for positioning purposes this errors should be compensated.

Camera-based position detection aided with accelerometer information was proposed by Tanaka et al. [42]. They included the rotation matrices resulting from the gravity vector estimation in the distance calculation equations. Accuracies below 5 cm are claimed within a 1.95-m-height scenario. The camera resolution is 1296 × 964 but no focal length or FOV is provided in the paper. The lamps used as reference are 0.5 m in diameter, do not transmit information, and each one is assigned a different color as ID.

Yang et al. [43] proposed a system that estimates position in the XY-plane using a single luminaire. It uses the typical configuration subject to the upwards-pointing restriction. The authors used a geometrical approximation without taking into account lens distortion. The experimental results claim an accuracy of around 2.5 cm. Nonetheless, the link range is not mentioned and relative error could not be extracted.

In [9], an IPS based on OCC using Undersampled Phase Shift Keying (UPSK) is experimentally evaluated. The approach followed in this work is the same as in the majority of related literature. The only contribution is the experimental evaluation of UPSK to transmit each lamp's ID. The positioning error reported is 5 cm at 120 cm.

Juneja and Vashisth described their Epsilon system [44]. In this camera-based VLP system, each LED bulb broadcasts location beacons to the receivers. The proposed beacons use Binary Frequency Shift Keying (BFSK). However, the authors only provided the theoretical framework.

Pan et al. proposed a camera-based visible light IPS with specular reflection cancellation [45]. The positioning algorithm is addressed and the geometrical relationship of LED beacons is used to identify and cancel the specular reflection. A mirror deployed close to a smartphone camera at a determined angle is used to create the specular reflection. The authors claimed that, using this method, more than 90% of the positioning error can be avoided, improving the robustness of the positioning system. The average error is 48.25 cm before cancellation and 1.53 cm after cancellation.

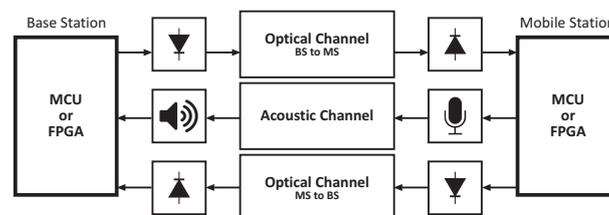
Zhang et al. [46] proposed a modification of the Single Value Decomposition (SVD) positioning algorithm in camera-based VLP. Generally, positioning algorithms estimate the orientation using homography when at least three LED emitters are located within the captured frame, and then solve the positioning algorithm as an optimization problem. The authors proposed a closed-form equation and compared their with the traditional iterative Levenberg–Marquardt algorithm, improving accuracy and convergence time.

Regarding surveys, some can be found in the literature. Concretely, in [47,48], some interesting issues are highlighted regarding VLP. In proximity positioning, an RF-based interface is yet needed to fulfill the location procedure. It is hoped that, in the future, VLC-based uplink can be used for this purpose. Outdoor VLC-based positioning is still a big challenge using traditional photo-receivers, and camera-based systems will be a hot research topic in the future. Finally, lens distortion is generally not addressed in the literature, and its capital to improve the positioning performance in OCC-based systems. Furthermore, motion blur in rolling-shutter cameras is generally neglected, but this highly affects both positioning and tracking, and must be studied more profoundly.

As can be observed from the presented state-of-the-art analysis, OCC-based positioning systems are based on the combination of lamp identification and geometrical relationships between the located emitters. In this work, a novel ranging technique based on the Rolling Shutter (RS) effect is presented. The proposed scheme uses the captured frames in an entirely different manner, allowing camera-based positioning algorithms combining the intrinsic angle estimation of image sensors and distance measurement.

### 3. System Description

As mentioned above, the proposed system is a variation of the scheme presented in [10]. Fundamentally, that system uses an optical signal instead of the RF one included in Cricket devices [49], and introduces an additional optical receiver and emitter to provide distance-measurement capability on both devices. The measurement procedure is triggered by the Base Station (BS), which sends both optical and ultrasound signals simultaneously. When the Mobile Station (MS) detects those signals, it carries out the distance estimation taking into account the time delay between signals. Besides, at the moment of receiving the ultrasound signal, the MS transmits an optical pulse. It is used as a reference for the distance estimation at the BS side. Distance is then calculated from the time lapse between the BS optical signal emission and the reception of the MS optical signal. Figure 1 depicts the block diagram of the system.



**Figure 1.** Hybrid Visible Light Communication-ultrasound system for distance estimation. Where MCU is a microcontroller unit and FPGA is a Field Programmable Gate Array.

Modifications proposed in this work mainly consist of substituting the BS optical receiver (photodiode) by a Complementary Metal Oxide Semiconductor (CMOS) camera, which implements a Rolling Shutter (RS) scanning method [50].

Photodiode-based receivers integrate all the incoming light (restricted by its attached optics FOV), generating an electrical signal that depends on the received power and its responsivity. This signal is typically conditioned using amplifiers and finally is analog-to-digital converted at a given sample rate, producing a continuous data stream. This type of receivers is characterized by its minimum detectable signal or sensitivity, and its bandwidth. When used in ranging applications, these parameters determine the maximum measurable range and the maximum distance resolution. Optical cameras can be considered as photodiode arrays attached to imaging optics and optical filters, in which each row is enabled sequentially. The scanning process limits the maximum received signal frequency and, hence, the maximum achievable time resolution. The maximum measurable range will depend on the minimum allowed image size for the performed ranging method. All these characteristics are studied in depth in the following sections, where Table 1 shows the maximum range and minimum resolution for different camera types.

In this new scheme, the time interval measurement is performed by an image processing algorithm instead of the classical pulse arrival detection used in TDoA systems. Furthermore, if audible elements replaced the ultrasound ones, the system's BS and MS would be suitable for a COTS devices-based implementation. The location process is thought to be performed, fully or partially, through commercial smartphones or webcams, microphones and speakers. That assumption makes cost implementation similar to a smartphone-based APP. If needed, some specific electronics can also be used in the Mobile Station, but no specific, ad-hoc hardware circuits are needed. Positioning systems require  $n + 1$  reference points on an  $n$ -dimensional space (e.g., three points to find position over a surface) to perform the trilateration scheme. In this case, three MS can be used as reference stations, reducing the cost. MS design leads to a very simple and cheap implementation, as only an optical and acoustic receiver, an optical emitter, and some basic control electronics are required, without further computation and communications elements. The new block scheme is shown in Figure 2.

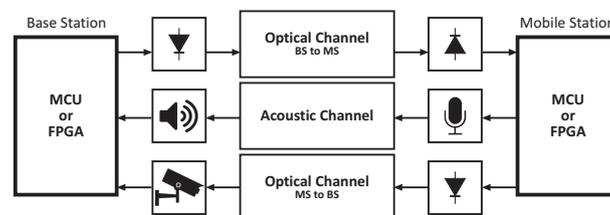
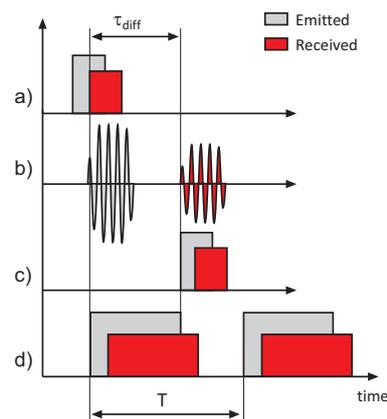


Figure 2. Block diagram of the proposed scheme.

In addition to the hardware modifications, some changes must be introduced in the signals generated in the measurement process. Chronograms for both systems are represented in Figure 3.

The ranging procedure attaining the MS-side distance estimation remains unchanged from the one in [10]. Nevertheless, when the MS receives the optical signal, it starts its optical emission and does not finish until the reception of the ultrasound signal. This is the main difference with the former system, in which the optical signal consists of a short pulse when the sound signal is detected. Therefore, the BS camera gets an optical signal whose duration is equal to the time difference of arrival at the MS.

To allow a camera-based detection, the MS does not emit a single pulse, but a pulse train. The period of this signal significantly affects the estimation error performance, as discussed below. Once the camera retrieves enough frames, the system will proceed to the distance estimation using image processing. This is described in the next section.



**Figure 3.** (a) Emitted and received optical signals from Base Station (BS) to Mobile Station (MS); (b) emitted and received ultrasound signals from BS to MS; (c) emitted and received optical pulse from MS to BS (former scheme, detected using a single optoelectronic device); and (d) emitted and received optical pulse train from MS to BS (new scheme, detected using a CMOS camera).

#### 4. System Performance and Constraints

TDoA-based ranging techniques are based on the different propagation speeds of two emitted signals. Generally, electromagnetic and pressure waves are used because of their extremely different propagation speeds. However, TDoA-based ranging could also be achieved using two different wavelengths whose corresponding refractive indices differ. The ranging procedure is started by the fastest signals and serves as the time reference. Indeed, in this work, the time to reach the receiver is small enough to neglect it in the calculations, as can be observed in the chronogram in Figure 3 (the transmitted and received optical signals present minimal delay). The slower one, generally a pressure wave such as ultrasound or acoustic emission, is the actual reference for the delay-difference measurement. Equation (1) shows how the distance estimation is carried out from these two signals.

$$\tilde{d} = \frac{v_1(T) \cdot v_2(T)}{v_1(T) - v_2(T)} \Delta t \approx_{v_1 \gg v_2} v_2(T) \cdot \Delta t \quad (1)$$

Where  $\tilde{d}$  is the estimated distance,  $v_1$  is the quicker signal speed,  $v_2$  is the slower signal speed, and  $\Delta t$  is the time interval between signals. Finally,  $T$  is temperature, which profoundly affects propagation speed, mostly in the case of pressure waves. Because of this, these systems also require a calibration based on temperature probing. The higher is  $v_1$  versus  $v_2$  the more accurate the distance estimation could be. In this work, an optical signal with  $v_1 = c_0$  speed is used, and a pressure wave (ultrasound or acoustic) is employed whose speed  $v_2$  is around 343 m/s, with some variation due to the ambient temperature. The measurement of  $\Delta t$  is usually performed employing a digital timer, which is triggered by these signals and its accuracy depends on the clock frequency of the electronics implementing the timer.

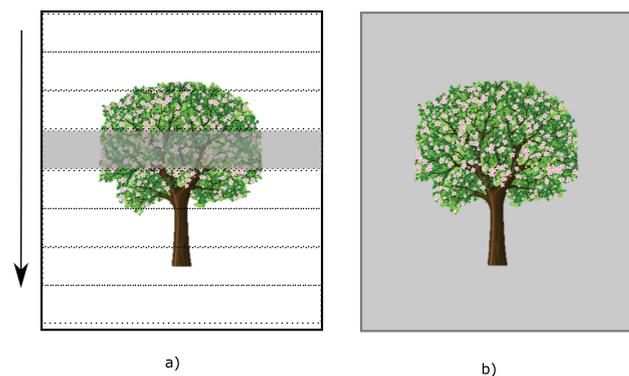
This work proposes two main modifications with respect to a conventional TDoA system: the optical signal detection device, and the time estimation procedure. Instead of a typical receiver for the optical signal such as a photodiode, a general purpose CMOS camera is used. The primary requirement is that the camera must be a rolling-shutter type one. The time estimation process is based on image processing techniques, instead of the conventional timer-based schemes.

The inclusion of a camera allows the use of a wide variety of COTS devices available in the market as a Personal Computer (PC) with a webcam, smartphones, tablets or any camera-equipped appliance (if not included, an LED and a speaker or ultrasound emitter should be integrated into the equipment). On the other hand, this kind of devices will presumably present worse time-measurement performance than the dedicated hardware timer of an embedded platform. Hence, a lower effective

time resolution is expected. Nevertheless, the ranging accuracy should suffice for most applications, as demonstrated below.

CCD (Charged Coupling Devices) and CMOS are both image sensors usually found in digital cameras, and they are responsible for converting light into electric signals. CCD are high-quality sensors that produce excellent images, but they are costly because they require unique manufacturing processes. CMOS sensors are much more inexpensive when compared to CCD since they can be manufactured on standard silicon integrated circuits production lines. That is the reason CMOS sensors have replaced CCD in the mass market and are being integrated into most cameras at present.

RS is a method of image capture used by CMOS cameras which does not expose the entire sensor simultaneously, but in a row basis. In Global Shutter (GS) systems, a usual technique found in CCD cameras, the whole sensor array integrates the incoming light simultaneously. RS technique activates the sensor array row by row from top to bottom of the picture (scanning across the scene rapidly), as shown in Figure 4. This difference produces predictable distortions of fast-moving objects (motion blur) or rapid flashes of light since the CMOS sensor top, and bottom parts capture different moments in time. This is a problem to be considered in image processing, but it is an advantage for OCC, since the scanning process allows capturing light changes with a higher time resolution than GS cameras (concretely sensor-height times greater).



**Figure 4.** Rolling Shutter vs. Global Shutter scanning methods. (a) rolling shutter; (b) global shutter.

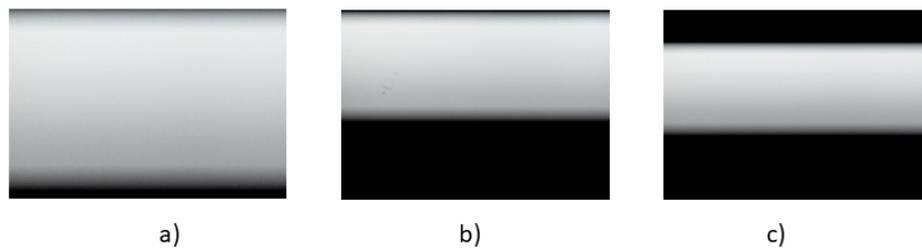
To obtain comparable results in both camera scanning types, the activation signal in RS devices needs to be faster than in GS. Therefore, considering a  $f_{frame}$  fps camera with a  $N_y$  rows sensor, a GS needs to activate its sensor  $f_{frame}$  times per second, while a RS element has to enable the rows for readout at a frequency of  $f_{frame} \cdot N_y$  rows  $\cdot$  s $^{-1}$ . This can be seen from the communications point of view as a different sampling frequency or a distinct time resolution for GS and RS devices, respectively. In this way, RS present an  $N_y$  times higher sampling rate than GS. Therefore, it can detect greater bandwidth signals or shorter pulses. The other main difference is that the received signal is distributed among the rows in RS devices, i.e., each line in the sensor takes a sample of the signal, while, for GS, the full sensor takes the sample.

On the other hand, the signal variation is visualized in different ways. GS shows variations of the illuminated area from one frame to other, whilst for RS, it is possible to appreciate partial illuminated areas instead of the whole light spot. It can be appreciated that the time duration is translated into space area, which is why the pulse duration  $T_{pulse}$  can be obtained from the illuminated area width. The portion of the illuminated area depends on the signal on-time duration, as Figure 5 shows. Short pulses excite a few rows, and wider pulses maintain the illumination for a greater

number of lines. Equation (2) calculates the number of illuminated rows  $n_{pulse}$  in the camera’s image for a camera with a frame speed of  $f_{frame}$  fps and a  $N_y$  rows vertical size image sensor.

$$n_{pulse} = f_{frame} \cdot N_y \cdot T_{pulse} \tag{2}$$

When pulse duration covers all the camera’s lines, the resulting signal corresponds to the system’s maximum measured distance. For example, a 60 fps and  $3.264 \times 2.448$  pixels (8 Mpixels) camera, such as that included in some smartphones, provides 2448 lines for an RS system. This corresponds to a sampling frequency of  $60 \times 2448 = 146.88$  kHz or, in other words, a time resolution of  $1/146,880 = 6.81 \mu s$ . As the system proposed here uses sound signals for distance measurement, and considering the speed of sound (343 m/s), the theoretical minimum distance change detected by this technique is  $\Delta d = 343 \times 6.81e^{-6} = 2.34$  mm. In the case of a GS system, due to the used scanning process, it is not possible to perform a similar distance estimation. In the same way, it is also possible to determine the maximum measured distance in RS, multiplying the number of lines of the image by  $\Delta d$ . For the previous example, the maximum reachable distance is  $d_{max} = 2.34e^{-3} \times 2448 = 5.73$  m. In Table 1, different video formats are presented determining the same parameters (sampling frequency, time resolution,  $d_{min}$ , and  $d_{max}$ ). As it can be appreciated from the example and Table 1, the proposed system provides enough resolution and range for a wide variety of applications relying on distance estimation.



**Figure 5.** Different illuminated areas corresponding to different distances. (a) short distance; (b) medium distance; (c) long distance.

**Table 1.** Resolution and range for different video formats.

Format	Resolution	Lines	Frame Rate (fps)	Sampling Frequency (kHz)	Time Resolution ( $\mu s$ )	$d_{min}$ (mm)	$d_{max}$ (m)
VGA	$640 \times 480$	480	25	12	83.33	28.58	13.72
		480	50	24	41.67	14.29	6.86
Standard Definition (SD)	$768 \times 576$	576	25	14.4	69.44	23.82	13.76
High Definition (HD)	$1280 \times 720$	720	25	18	55.56	19.09	13.72
		720	50	36	27.78	9.53	6.86
Full HD	$1920 \times 1080 p$	1080	25	27	37.04	12.70	13.72
Ultra HD	$3840 \times 2160$	2160	25	54	18.52	6.35	13.72
Full Dome	$4096 \times 4096$	4096	30	122.88	8.14	2.79	11.48

However, in both cases (RS and GS), the transmitted information is contained within the projection of the light source on the image sensor. The lamp’s spot size in the captured image depends on the actual lamp’s size and the distance between lamp and camera. Figure 6 depicts the geometric diagram.

$d$  is the distance between lamp and camera,  $H$  and  $A$  are the actual lamp’s dimensions and  $D_y$  and  $D_x$  are the vertical and horizontal lamp’s dimensions in the camera’s image (in pixels).  $N_y$  and  $N_x$  are the total image’s size in pixels.  $\Theta_{LY}$  and  $\Theta_{LX}$  are the angles from the camera position corresponding

to the vertical and horizontal edges of the lamp, and  $FOV_v$  and  $FOV_h$  are the camera’s vertical and horizontal Fields of View, which determine the limits of the area scanned by the camera. Therefore, in the case of RS, the amount of data that can be transmitted within each frame time depends on the source’s height in the image. Concretely, to make the data retrievable by the OCC receiver, the maximum number of rows the transmitted packet can occupy on the picture must satisfy the restriction of Equation (3).

$$f_{frame} \cdot N_y \cdot T_{packet} \leq \frac{D_y}{2} \tag{3}$$

where  $T_{packet}$  is the packet duration. Furthermore, since there is no synchronization between the image sensor and the light source, the receiver must capture at least two frames per packet and, hence, the emitter must repeat its transmission at least once during that two frames. In the case of the proposed ranging system, the latter restriction can be reformulated since the “packet” is symmetrical and only comprises the pulse duration and a guard time  $T_{guard}$ , which separates the transmitted pulses. Since the objective of the measurement system is to estimate the pulse duration (which is directly related to distance), the lamp’s image size must allocate at least two pulses and one guard time for error-free measurement, yielding to Equation (4). The value of  $T_{guard}$  must ensure that the packet is entirely allocated within the lamp’s projection. Furthermore, this guard time must be long enough to allow correct pulse duration estimation. High  $T_{guard}$  values allow easy pulse discrimination but reduce the error-free measurement range since the packet would not fit into the source’s image for longer distances.

$$\begin{aligned} f_{frame} \cdot N_y \cdot (2T_{pulse} + T_{guard}) &\leq D_y + 1 \\ 2n_{pulse} + n_{guard} &\leq D_y + 1 \end{aligned} \tag{4}$$

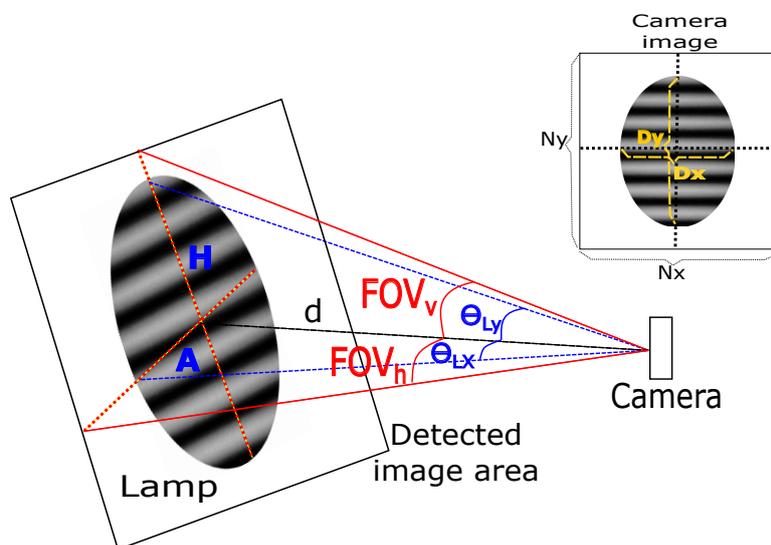


Figure 6. Geometrical scheme for spot size calculation.

As mentioned above, the light source’s projected size plays a capital role on the system’s performance. Concretely,  $D_y$  follows Equation (5) [51].

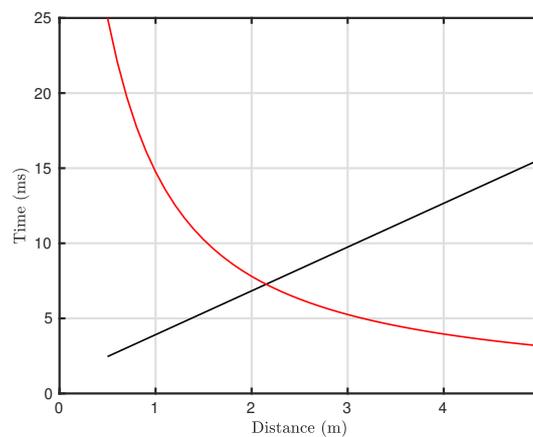
$$D_y = \frac{2N_y}{FOV_v} \arctan \left( \frac{H}{2d} \right) \tag{5}$$

where  $H$  is the apparent physical height of the radiant source,  $FOV_v$  is the receiver’s vertical field of view, and  $d$  is the actual separation distance between transmitter (lamp) and receiver (camera).

All these elements are depicted in Figure 6. Assuming the BS-to-MS ranging is carried out without errors,  $T_{pulse}$  can be defined as  $T_{pulse} = \gamma \cdot d/v_2(T)$ .  $\gamma$  is a scaling factor introduced by the authors as a way of improving the system’s performance. It modifies the pulse duration at MS-side to allow measurements at longer distances or to increase accuracy. In long-distance mode ( $\gamma < 1$ ), the generated pulse is shorter than the corresponding to the actual distance. In this way, whole pulses can fit into the resulting smaller Region Of Interest (ROI) due to distance. For instance,  $\gamma = 1/2$  would double the maximum measurement range. As a restriction, the receiver must be aware of this mode in order to compensate the effect of  $\gamma$  and to calculate the actual pulse duration. On the other hand, when in short-distance mode ( $\gamma > 1$ ), the generated pulse is wider than the expected one. Thus, very short distances can be measured because the pulse will cover enough lines in the image. As an example,  $\gamma = 2$  will double the resolution and the minimum-reachable distance to the half. Again, the receiver obtains the actual pulse duration compensating the  $\gamma$  factor. Introducing this into Equations (5) and (4) yields a new distance-dependent restriction (Equation (6)).

$$f_{frame} \cdot N_y \cdot \left( 2\gamma \cdot \frac{d}{v_2(T)} + T_{guard} \right) \leq \frac{2N_y}{FOV_v} \arctan \left( \frac{H}{2d} \right) + 1 \tag{6}$$

Figure 7 depicts the critical distance at which the condition of Equation (6) is violated. Beyond this distance, a correct recovery of the pulse duration (as a difference between rising and falling edges) is not ensured, and several runs of the measurement procedure (capturing more frames) must be carried out and statistical treatment must be applied. Typically, error uncertainty is reduced by averaging. However, for the presented scenario, this approach does not offer the best performance, as discussed in Section 5.



**Figure 7.** Critical distance shown as the intersection between the pulse period (on time plus guard, black curve) and the maximum allowed period given the object’s projected height (red curve). The curves correspond to both sides of the inequality in Equation (6) divided by  $f_{frame} \cdot N_y$  resulting in time units.  $\gamma = 1$ ,  $T_{guard} = 1$  ms,  $H = 1$  m,  $f_{frame} = 60$  fps, and  $FOV_v = \pi/3$  rad.

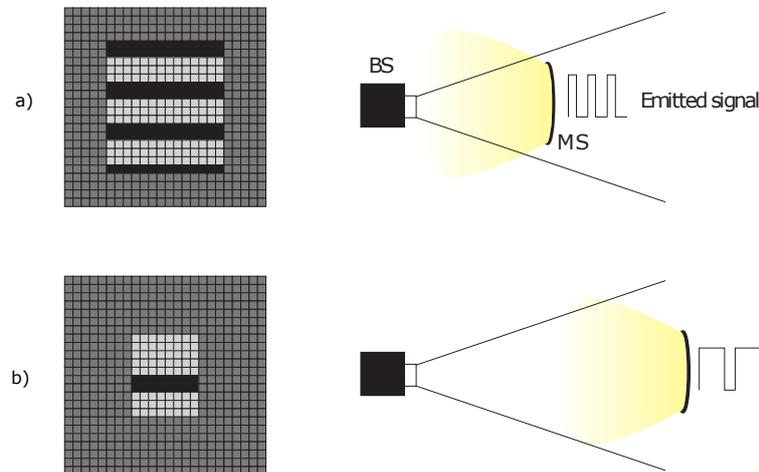
### 5. Statistical Analysis of the OCC Subsystem

As mentioned in Section 3, error-less distance estimation could be achieved if Equation (6) were satisfied. However, there still exists an uncertainty given by the time resolution  $\delta t$  of the RS sensor. This resolution corresponds to the row-sweeping time and is presented in Equation (7).

$$\delta t = \left( f_{frame} \cdot N_y \right)^{-1} \tag{7}$$

Therefore, all distance measurements are subject to an uncertainty of  $v_2(T) \cdot \delta t$ . Hereinafter, for simplicity but without loss of generality, this uncertainty is neglected. Furthermore, the pulse

duration is calculated as the maximum size of all the detected pulses within the Region Of Interest (ROI) corresponding to the light source. Figure 8 illustrates the procedure on a scenario in which several pulses are allocated, and another in which there is not enough space to detect a complete pulse. It must be taken into account that, if the ROI were smaller than  $n_{pulse}$ , it would be impossible to estimate distance properly, establishing a measurable-distance upper bound.



**Figure 8.** Example of detected OCC signal. (a) A closer object would blink at a higher frequency, and its projected image would be large. (b) A further object would be projected in a smaller ROI and the period of the transmitted signal would increase.

In this section, a statistical analysis of the distance estimation error is presented. Furthermore, the implications of the proposed error-mitigation technique on the measurement time are also analyzed. First, let us consider  $M$  runs of the presented pulse detection algorithm. For each measurement, the illuminated row count  $n$  (or detected pulse duration) would follow a probability density function (pdf) determined by  $D_y$ ,  $n_{pulse}$ , and  $n_{guard}$ . This pdf  $f_n(n)$  is discrete and is governed by Equation (8). It was assumed that the starting point of the pulsed signal follows a uniform distribution. The pdf was calculated analyzing the amount of cases for each number of rows  $n$  within a determined window of height  $D_y$ .

$$f_n(n) = \begin{cases} \frac{D_y - n_{pulse} + 1}{n_{pulse} + n_{guard}} & n = n_{pulse} \\ \frac{2}{n_{pulse} + n_{guard}} & n = n_{pulse} - i \quad / i \in [1, \lfloor n_{pulse}/2 \rfloor] \\ \frac{1}{n_{pulse} + n_{guard}} & n = n_{pulse}/2 \quad / D_y - n_{guard} \text{ even} \end{cases} \quad (8)$$

where  $\lfloor \cdot \rfloor$  is the floor operation. Note that the pdf can be roughly approximated by a Dirac's delta at  $n_{pulse}$  plus a uniformly distributed function at the rest of possibilities. After this pdf, a row-count error pdf  $f_r(r)$  could be derived from the difference between the correct detection ( $n_{pulse}$ ) and the performed measurement (Equation (9)). In this case, the change of variable  $r = n_{pulse} - n$  was applied and the pdf from Equation (8) appropriately modified.

$$f_r(r) = \begin{cases} \frac{D_y - n_{pulse} + 1}{n_{pulse} + n_{guard}} & r = 0 \\ \frac{2}{n_{pulse} + n_{guard}} & r = i \quad / i \in [1, \lfloor n_{pulse}/2 \rfloor] \\ \frac{1}{n_{pulse} + n_{guard}} & r = n_{pulse}/2 \quad / D_y - n_{guard} \text{ even} \end{cases} \quad (9)$$

This error can be scaled to distance error using Equations (1) and (7), yielding the final distance-measurement error pdf  $f_e(e)$  (Equation (10)). This error is the difference between the actual distance ( $d$ ) and the estimated one ( $\hat{d}$ ). The estimated distance depends on the number of detected rows ( $n$ ), the sound speed ( $v_2(T)$ ), and the time resolution ( $\delta t$ ).

$$f_e(e) = \begin{cases} \frac{D_y - n_{pulse} + 1}{n_{pulse} + n_{guard}} & e = 0 \\ \frac{2}{n_{pulse} + n_{guard}} & e = i \cdot v_2(T)\delta t \\ \frac{1}{n_{pulse} + n_{guard}} & e = n_{pulse}/2 \cdot v_2(T)\delta t \end{cases} \quad \begin{matrix} /i \in [1, \lfloor n_{pulse}/2 \rfloor] \\ /D_y - n_{guard} \text{ even} \end{matrix} \quad (10)$$

The three probability density functions present the same shape, since their underlying variables ( $n$ ,  $r$ , and  $e$ ) are related by linear transformations. This means that the final distance error depends on the row-counting mismatch. When the light source’s projection is large enough, the distance error measurement would tend to zero since the pulsed signals would be entirely allocated within the ROI. On the other hand, as the ROI decreases (which is the scenario corresponding to the presented pdfs), the detected number of activated rows could be lower than the actual period of the signal ( $n_{pulse}$ ). When this occurs, there are three possibilities:

- The pulse is allocated entirely ( $n = n_{pulse}$ ,  $r = 0$ ,  $e = 0$ ). This corresponds to zero error.
- The detected pulse size is smaller than the actual one ( $n = n_{pulse} - i$ ,  $r = i$ ,  $e = i \cdot v_2(T)\delta t$ ). This part of the pdf corresponds to a uniform distribution, whose span ranges from 1 to  $n_{pulse}/2$ .
- The last item corresponds to the case in which the detected pulse is half the actual one. This only occurs when the difference between  $D_y$  and  $n_{pulse}$  is even.

Since the shape of  $f_e(e)$  is asymmetrical, error reduction based on averaging is not optimal. However, there is a predominance of the zero-error measurement with respect to the uniformly distributed rest. In this work, an error-reduction technique based on the mode of a sequence of  $M$  measurements ( $\vec{n}$ ) is proposed (Equation (11)).

$$\tilde{n}_{pulse} = \sup \text{mode}(\vec{n}) \quad (11)$$

where  $\vec{n} = \{n_0, n_1, \dots, n_{M-1}\}$  is a  $M$ -element vector of measurements,  $\tilde{n}_{pulse}$  is the final estimated pulse length, and  $\sup \text{mode}(\cdot)$  is the maximum value of the mode (since several maxima can be found on  $\vec{n}$ ). Focusing (for simplicity) on a situation in which  $D_y - n_{guard}$  were odd, and naming  $p$  and  $q$  the probabilities of no measurement error and the rest respectively, the following analysis regarding  $\vec{n}$  can be made.

$$\begin{aligned} \# \{ \vec{n} = n_{pulse} \} &\sim B(M, p) \\ \# \{ \vec{n} = n_{pulse} - i \} &\sim B\left(M - \sum_{j=0}^{i-1} \# \{ \vec{n} = n_{pulse} - j \}, q\right), \\ \sum \# \{ \vec{n} = n_i \} &= M \end{aligned} \quad (12)$$

where  $\#\{\cdot\}$  denotes cardinality, and  $B(M, p)$  is the Binomial distribution. The last equations establish the statistical behavior of the absolute frequencies for each possible measurement. The number of correct measurements ( $n = n_{pulse}$ ) follows a Binomial distribution with  $M$  tries with success probability  $p$  (first equation).  $p$  corresponds to  $f_n(n_{pulse})$  from Equation (8). The second equation describes the number of occurrences for each row count. It must be taken into account that this definition arises from conditional probabilities, and the Binomial distribution of the  $j$ th element depends on both the boundary condition of the third equation and the distribution of the  $(j - 1)$ th element. This definition leads to a  $M - 1$  dimensional probability density function, in which each capture is considered independent. In this case, the mode is associated with the superior argument maximum of the cardinality vector. This problem must be solved numerically, which is illustrated in Section 6.

### 6. Simulation Results

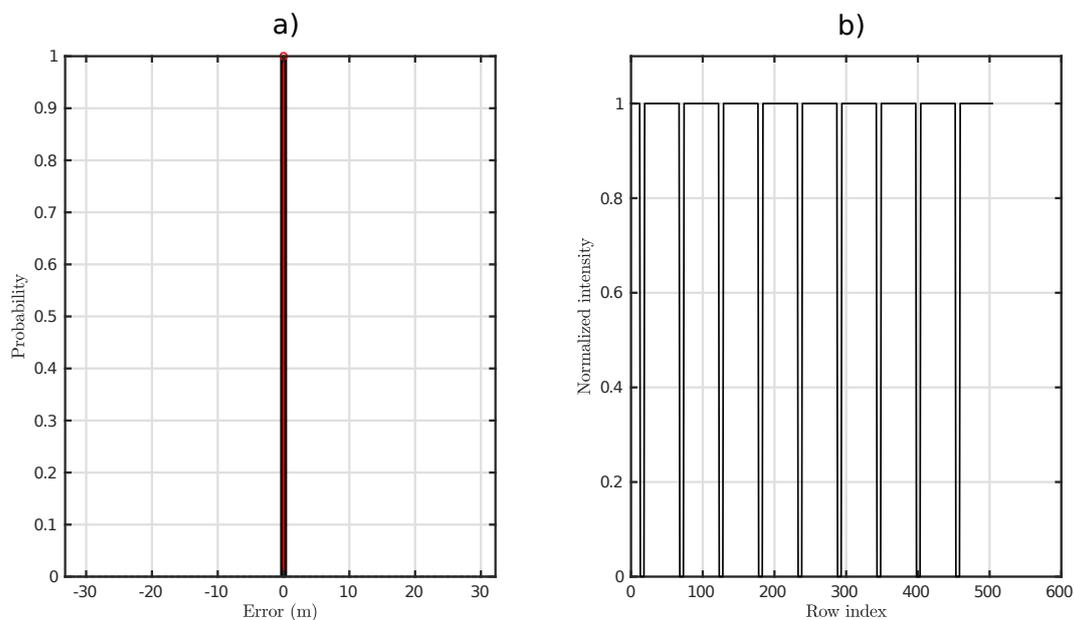
This section presents several results regarding the system performance concerning error. To manifest the benefits of basing the estimation on the statistical mode instead of the mean value, both techniques were compared. The histograms associated with the averaging strategy versus the mode-based proposal is depicted, highlighting and justifying the performance improvement.

As a baseline, the parameters in Table 2 were used to obtain the simulation results. Variations on some parameters were analyzed to evaluate their impact on the system’s performance.

**Table 2.** Baseline parameters of the simulation.

Parameter	Value
$FOV_v$	60°
$N_y$	1080 rows
$f_{frame}$	60 fps
$H$	50 cm
$v_2(T)$	333 m·s <sup>-1</sup>
$T_{guard}$	100 μs

Figure 9 depicts the simulated measurement error histogram at 1 m versus the theoretical curve extracted from Equation (10). In the same figure, it can be observed that there are no errors since several pulses are allocated within the ROI. As mentioned above, the pulse width depends linearly on distance (Equation (2)). In addition, due to the proximity of the measurement, the lamp’s projection would be large enough to allocate several pulses.

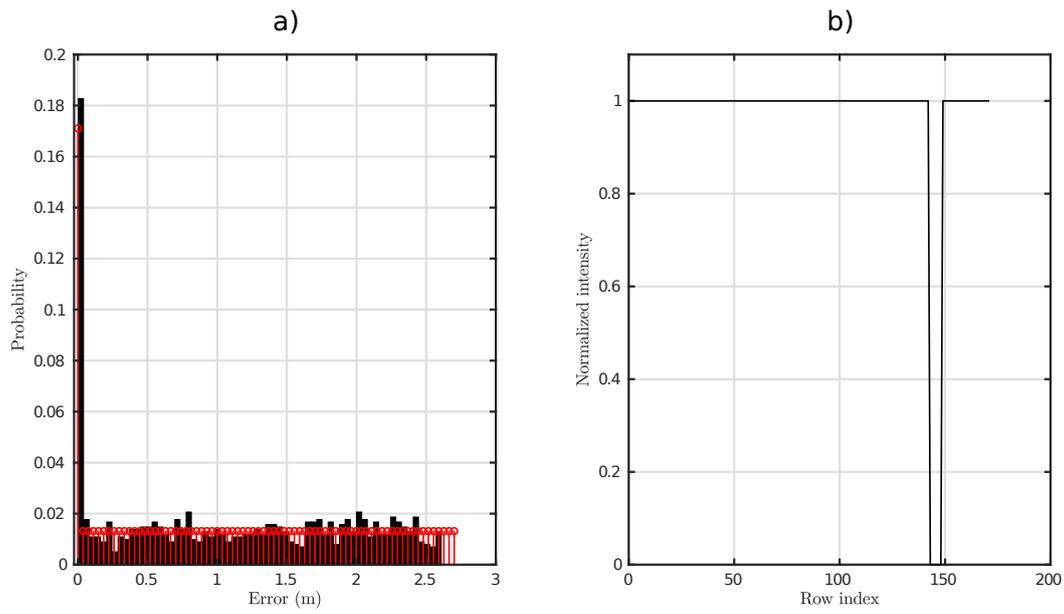


**Figure 9.** Empirical histogram in black versus theoretical result in red (a); and simulated signal within the corresponding ROI at 1 m (b). In total, 1024 independent simulations were performed.

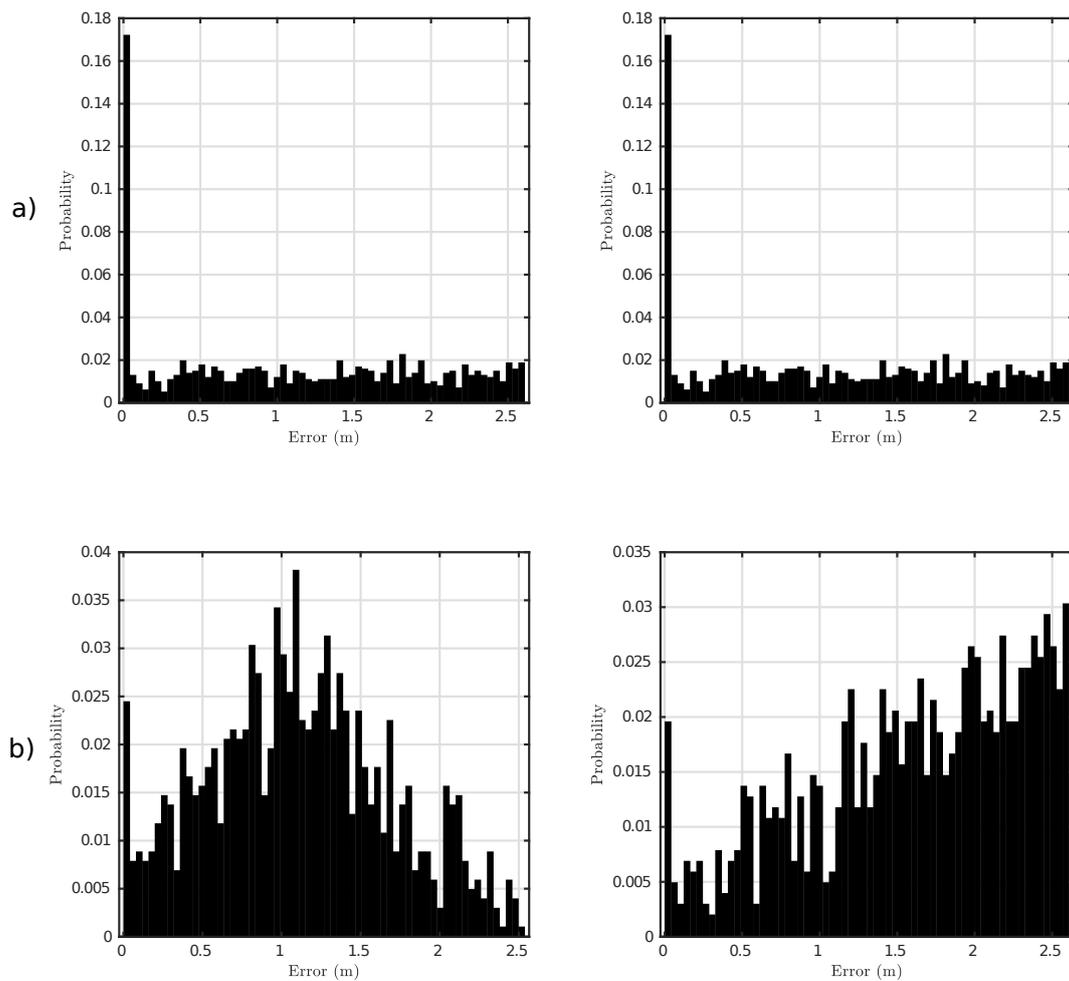
Locating emitter and receiver at a distance above the critical distance defined by Equation (6), other components appear in the histogram, as Figure 10 depicts. In this case, the lamp’s projection has been reduced and the pulse duration has been increased due to distance. Since the scenario is not compliant with the restriction of Equation (4), the probability of allocating an entire pulse is reduced. Thus, erroneous measurements may occur.

As discussed in the previous sections, since the pdf is asymmetric and the zero-error component corresponds to the most frequent value of the distribution, the mode is proposed as an error-mitigation strategy. Figure 11 depicts the histograms of mean and mode for a different number of measures at the distance shown in Figure 10.

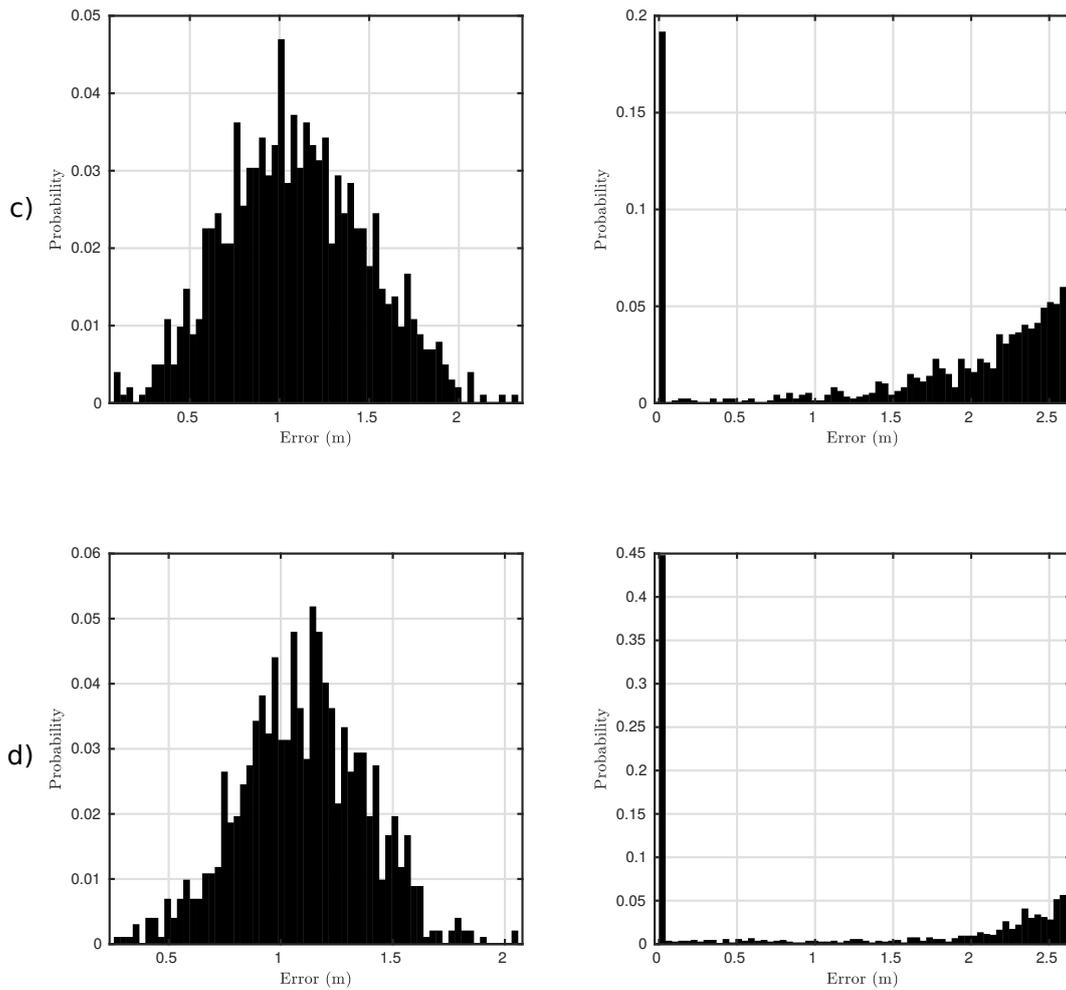
Figure 12 shows the evolution of the expected value of the error as a function of distance and number of runs for both error-mitigation strategies. Figure 13 depicts in detail the part of Figure 12 ranging from 2 to 3 m.



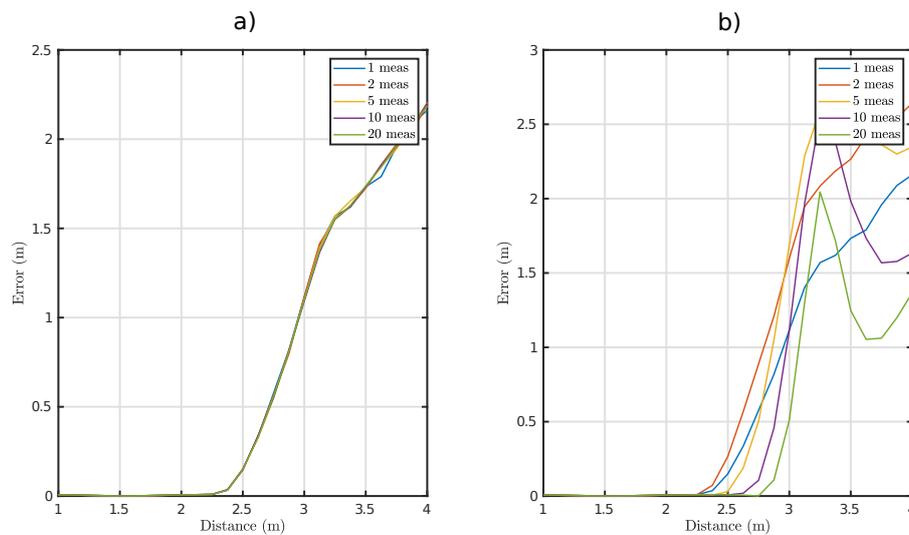
**Figure 10.** Empirical histogram in black versus theoretical result in red (a); and simulated signal within the corresponding ROI at 3 m (b). In total, 1024 independent simulations were performed.



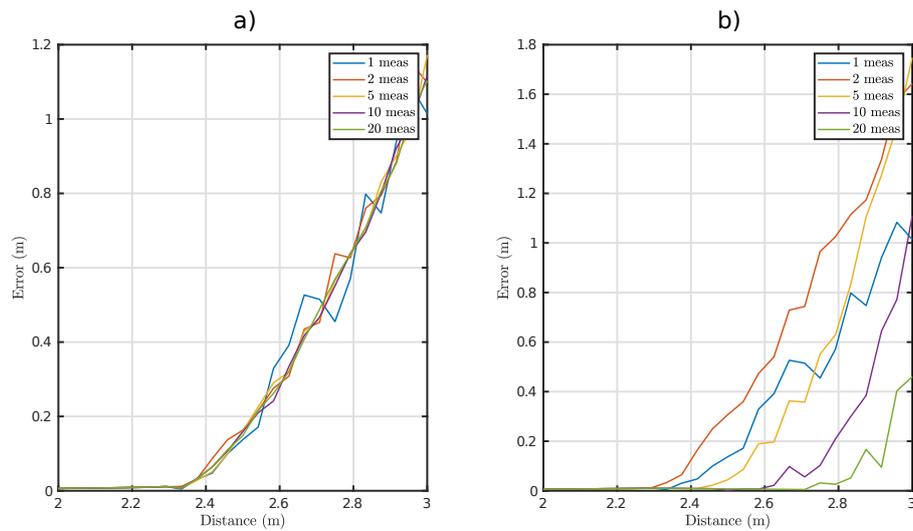
**Figure 11.** Cont.



**Figure 11.** Histograms associated to the averaging strategy (**left column**); and to the proposed mode-based error mitigation (**right column**). (**a**) one-shot measurement samples filtering, (**b**) two-shot measurements samples filtering, (**c**) five-shot measurements samples filtering and (**d**) ten-shot measurements samples filtering.

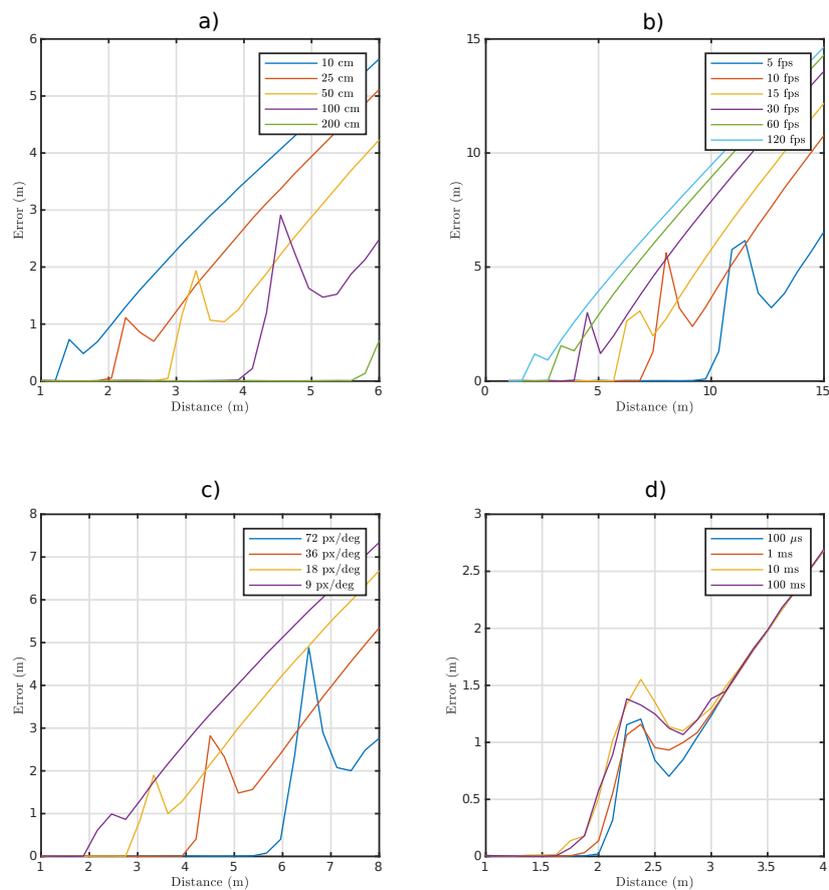


**Figure 12.** Expected error versus distance for different number of runs: (**a**) the typical averaging strategy; and (**b**) the proposed mode-based filtering.



**Figure 13.** Zoom of the error performance in Figure 12. (a) the typical averaging strategy; and (b) the proposed mode-based filtering.

Finally, Figure 14 illustrates how the variation of the object’s effective size, capture rate, guard time and FOV-to-resolution ratio influence the error performance.



**Figure 14.** Effect of different parameters on error performance: object’s effective size (a); capture rate (b); FOV-to-resolution ratio (c); and guard time (d). All curves were obtained using the proposed mode-based filtering with 20 samples.

## 7. Discussion

In this work, a new distance estimation technique, based on TDoA and VLC OCC systems, is presented. It consists of a variation of another proposal of the authors, which uses optical pulses instead of RF signal as in Cricket systems. The novelty of the present system is the use of a camera as an optical receiver and the performance of time intervals calculations employing image processing techniques. The proposed method is based on the characteristic of the RS cameras, which allows the association of a pulse time interval to the height of the illuminated area produced by this pulse in the camera's captured picture.

The system's resolution and range depend on the camera's video recording frames per seconds and the image's number of lines used in the RS CMOS sensor. Some millimeters of resolution and tens of meters range are viable capabilities of this kind of systems, and this will be improved with the introduction of faster (more fps) and higher quality (more RS lines) cameras. Furthermore, some advantages are introduced by the proposed scheme such as the capability of several simultaneous distance measurements, since several spots can be present in the same image, and the possibility of being integrated into commercial devices. Smartphones, tablets, and other appliances are provided with LEDs (flash, for example), sound emitters (loudspeaker) and receivers (microphone), and cameras, which are the essential components of the devices used in the proposed system.

The results presented in the previous section assume no error in the hybrid optoacoustic TDoA ranging stage. In actual implementations of this system, the measured time difference would be affected directly by the detection electronics of both optical and ultrasound front-ends. In addition, the BS's emitted power and the channel's path loss would also impair error performance. Nonetheless, taking into account that the OCC-based feedback limits the maximum achievable ranging distance, it is not unreasonable to assume at least centimeter accuracies.

From the extracted mathematical formulation, there is a range of distances in which the error would be uniquely subject to the uncertainty defined by the row-count rate  $\delta t$ . However, beyond the above-defined critical distance, errors induced by the appearance of partial pulses in the ROI may appear. This distance depends on both geometric and receiver parameters, as can be observed in Figure 14.

The histograms in Figures 9 and 10 validate the mathematical formulation extracted in Section 5. Furthermore, it is straightforward to demonstrate that averaging on a set of independent measurements does not eliminate the error but introduces a distance-dependent bias.

From the shape of the error pdf  $f_e(e)$ , the use of the superior mode as an error-mitigation technique is proposed. The improvement on the error performance can be observed in Figure 12. In the case of the traditional averaging strategy, a higher number of runs does not offer an advantage regarding expected value. However, from the histograms presented in Figure 11, a smaller variance is expected from the averaging. Regarding the mode-base technique, due to the intrinsic nonlinearity of the statistical mode and the shape of the pdf, both expected value and variance are dramatically reduced. Observing Figure 12, a significant improvement of the critical distance can be obtained using just 20 measurements, which is translated to acquiring and processing 20 frames.

Variations on several parameters were simulated and the results can be observed in Figure 14. As expected, the bigger the object is, the longer the measured distance can be. This relation is nonlinear and is governed primarily by Equation (6). In the case of the camera's capture rate, this relation is inverse. This occurs because the slower the row-reading rate, the smaller an object's projection on the image sensor can be. However, the uncertainty of the measure would also increase. The angular accuracy (vertical resolution over FOV) has the same impact as the object's physical size since it affects the projected image in the same manner. Finally, the smaller  $T_{guard}$  is, the more possible acquiring complete pulses within the ROI is (Equation (6)). Furthermore, significant values of  $T_{guard}$  may lead to a situation in which some frames could present energy-less ROI. This means that no pulse would be allocated inside and, hence, more frames must be acquired with the subsequent increment of the overall measurement time.

Table 3 compares the system's error performance and complexity with the current state-of-the-art image-based location systems. It can be observed that the use of the proposed TDoA-assisted OCC-based ranging enhances the performance compared to traditional image-based systems. This table shows the performance using the baseline parameters of Table 2, but they could be pushed further by properly selecting them.

**Table 3.** Performance comparison between the proposed system and other works available in the literature.

System	Error	Technique	Complexity
Lin et al. [9]	5 cm @ 1.2 m	OCC	Very low
Li et al. [14]	6 cm @ 0.8 m	OCC	Very Low
Hossar et al. [15]	10 cm @ 3 m	OCC	Low
Tanaka et al. [42]	5 cm @ 1.95 m	OCC	Low
Chen et al. [22]	>10 cm @ 1 m	VLC	Low
Chen and Lain [23]	3 cm @ 1 m	OCC	Low
Alam et al. [32]	2.2 cm @ 2.4 m	Fingerprint	High
Konings et al. [33]	10 cm @ 2.4 m	VLC + RF Fingerprint	High
Proposed scheme	sub-cm @ 2.2 m	OCC + TDoA	Low
Proposed scheme with mode	sub-cm @ 2.7 m	OCC + TDoA	Low

In this work, blooming has been neglected assuming that the irradiance is small enough, but it would affect the error introducing a distance-independent bias and lower bound on  $T_{guard}$ . Therefore, the potential effect of blooming could be easily calibrated and compensated.

A possible limitation of the proposed strategy is the possibility of synchronism between emitter and receiver. This occurs when the emitter's blinking rate is an integer (or close) multiple of the receiver's capture rate. Then, the acquired frames would not comply with the assumed statistical uncorrelation. To mitigate this harmful effect, a random back-off (e.g., exponential back-off) could be used instead of a fixed  $T_{guard}$ .

Finally, the effect of background illumination (sunlight or other light sources) has not been analyzed in this work. Nevertheless, its effects on the measurement performance would be presumably negligible. On the one side, the use of imaging optics maps solid angles to camera pixels, separating the measurement lamp and other illumination sources in the captured frames. This reduces to zero the interference unless the other background sources were overlapped with the ROI. On the other hand, reflections of these background sources (including the sunlight) on the measurement lamp could introduce bias to the perceived signal (the lamp's projection would present some energy even when is OFF). This effect could reduce the peak-to-peak value of the received signal. Nonetheless, these contributions are first- or second-order reflections, which imply very low energy, and hence very small impact on the system's performance.

**Author Contributions:** Conceptualization, J.R. (Jose Rabadan) and V.G.; Data curation, V.G. and J.R. (Julio Rufo); Formal analysis, J.R. (Jose Rabadan), V.G., C.G., J.R. (Julio Rufo) and R.P.-J.; Funding acquisition, J.R. (Jose Rabadan) and R.P.-J.; Investigation, J.R. (Jose Rabadan) and V.G.; Methodology, J.R. (Jose Rabadan); Project administration, J.R. (Jose Rabadan); Resources, R.P.-J.; Software, V.G. and C.G.; Supervision, J.R. (Jose Rabadan); Visualization, V.G., C.G. and J.R. (Julio Rufo); Writing—original draft, J.R. (Jose Rabadan), V.G. and C.G.; and Writing—review and editing, J.R. (Jose Rabadan), V.G., C.G., J.R. (Julio Rufo) and R.P.-J.

**Funding:** This research was funded in part by the Spanish Research Administration (MUCIU, Project OSCAR, Ref. TEC2017-84065-C3-1-R) and the Canary Islands Regional Government (Project ATICCuA ProID2017010053)

**Conflicts of Interest:** The authors declare that they do not have any conflict of interest regarding the contents of this paper.

## References

1. Waadt, A.; Bruck, G.; Jung, P. *Mobile Positioning and Tracking: From Conventional to Cooperative Techniques*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2010; pp. 177–211. [[CrossRef](#)]
2. Correa, A.; Barcelo, M.; Morell, A.; Vicario, J.L. A Review of Pedestrian Indoor Positioning Systems for Mass Market Applications. *Sensors* **2017**, *17*, 1927. [[CrossRef](#)] [[PubMed](#)]

3. Perez-Jimenez, R.; Rabadan, J.; Guerr, V.; Aguiar, L.; Rufo, J. Fundamentals of Indoor VLP: Providing Autonomous Mobility for Visually Impaired People. In Proceedings of the 2017 IEEE International Conference and Workshop on Bioinspired Intelligence (IWOBI), Funchal, Portugal, 10–12 July 2017; pp. 1–6. [\[CrossRef\]](#)
4. Armstrong, J.; Ahmet Sekercioglu, Y.; Neild, A. Visible light positioning: A roadmap for international standardization. *IEEE Commun. Mag.* **2013**, *51*, 68–73. [\[CrossRef\]](#)
5. Iturralde, D.; Azurdia-Meza, C.; Krommenacker, N.; Soto, I.; Ghassemlooy, Z.; Becerra, B. A new location system for an underground mining environment using visible light communications. In Proceedings of the 2014 IEEE 9th International Symposium on Communication Systems, Networks & Digital Sign (CSNDSP), Manchester, UK, 23–25 July 2014; pp. 1165–1169. [\[CrossRef\]](#)
6. Munoz, D.; Lara, F.B.; Vargas, C.; Enriquez-Caldera, R. *Position Location Techniques and Applications*; Academic Press, Inc.: Orlando, FL, USA, 2009.
7. Xu, X.; Tang, Y.; Li, S. Indoor localization based on hybrid Wi-Fi hotspots. In Proceedings of the 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sapporo, Japan, 18–21 September 2017, pp. 1–8. [\[CrossRef\]](#)
8. Zhang, W.; Kavehrad, M. Comparison of VLC-based indoor positioning techniques. In Proceedings Broadband Access Communication Technologies, San Francisco, CA, USA, 4 February 2013; Volume 8645, p. 86450M. [\[CrossRef\]](#)
9. Lin, B.; Ghassemlooy, Z.; Lin, C.; Tang, X.; Li, Y.; Zhang, S. An Indoor Visible Light Positioning System Based on Optical Camera Communications. *IEEE Photonics Technol. Lett.* **2017**, *29*, 579–582. [\[CrossRef\]](#)
10. Rabadan, J.; Guerra, V.; Rodríguez, R.; Rufo, J.; Luna-Rivera, M.; Perez-Jimenez, R. Hybrid Visible Light and Ultrasound-Based Sensor for Distance Estimation. *Sensors* **2017**, *17*, 330. [\[CrossRef\]](#)
11. Wang, C.; Wang, L.; Chi, X.; Liu, S.; Shi, W.; Deng, J. The research of indoor positioning based on visible light communication. *China Commun.* **2015**, *12*, 85–92. [\[CrossRef\]](#)
12. Fang, J.; Yang, Z.; Long, S.; Wu, Z.; Zhao, X.; Liang, F.; Jiang, Z.L.; Chen, Z. High-Speed Indoor Navigation System based on Visible Light and Mobile Phone. *IEEE Photonics J.* **2017**, *9*, 1–11. [\[CrossRef\]](#)
13. Lin, B.; Tang, X.; Li, Y.; Zhang, M.; Lin, C.; Ghassemlooy, Z.; Wei, Y.; Wu, Y.; Li, H. Experimental demonstration of optical camera communications based indoor visible light positioning system. In Proceedings of the 2017 16th International Conference on Optical Communications and Networks (ICOON), Wuzhen, China, 7–10 August 2017; pp. 1–3. [\[CrossRef\]](#)
14. Li, Y.; Ghassemlooy, Z.; Tang, X.; Lin, B.; Zhang, Y. A VLC Smartphone Camera Based Indoor Positioning System. *IEEE Photonics Technol. Lett.* **2018**, *30*, 1171–1174. [\[CrossRef\]](#)
15. Hossan, T.; Chowdhury, M.Z.; Islam, A.; Jang, Y.M. A Novel Indoor Mobile Localization System Based on Optical Camera Communication. *Hindawi Wirel. Commun. Mob. Comput.* **2018**, *2018*, 9353428. [\[CrossRef\]](#)
16. Chavez-Burbano, P.; Guerra, V.; Rabadan, J.; Jurado-Verdu, C.; Perez-Jimenez, R. Novel Indoor Localization System Using Optical Camera Communication. In Proceedings of the 2018 11th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP), Budapest, Hungary, 18–20 July 2018; pp. 1–5. [\[CrossRef\]](#)
17. Dong, Z.; Lu, F.; Ma, R.; Wang, L.; Zhang, C.; Chen, G.; Wang, A.; Zhao, B. An integrated transmitter for LED-based visible light communication and positioning system in a 180 nm BCD technology. In Proceedings of the 2014 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM), Coronado, CA, USA, 28 September–1 October 2014; pp. 84–87. [\[CrossRef\]](#)
18. Lou, P.; Zhang, H.; Zhang, X.; Yao, M.; Xu, Z. Fundamental analysis for indoor visible light positioning system. In Proceedings of the 2012 1st IEEE International Conference on Communications in China Workshops (ICCC), Beijing, China, 15–17 August 2012; pp. 59–63. [\[CrossRef\]](#)
19. Do, T.H.; Hwang, J.; Yoo, M. TDoA based indoor visible light positioning systems. In Proceedings of the 2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN), Da Nang, Vietnam, 2–5 July 2013; pp. 456–458. [\[CrossRef\]](#)
20. Roberts, R.; Gopalakrishnan, P.; Rathi, S. Visible light positioning: Automotive use case. In Proceedings of the 2010 IEEE Vehicular Networking Conference, Jersey City, NJ, USA, 13–15 December 2010; pp. 309–314. [\[CrossRef\]](#)

21. Bai, B.; Chen, G.; Xu, Z.; Fan, Y. Visible Light Positioning Based on LED Traffic Light and Photodiode. In Proceedings of the 2011 IEEE Vehicular Technology Conference (VTC Fall), San Francisco, CA, USA, 5–8 September 2011; pp. 1–5. [[CrossRef](#)]
22. Chen, C.; He, Z.; Huang, K.; Hu, J.; Liu, B.; Huang, Q.; Huang, X.; Liu, Q.; Jiang, C.; Peng, Q.; et al. Visible light communication using the microphone jack of the smart phone as an optical receiver and its application in the indoor localization system. In Proceedings of the 2017 International Conference on Electron Devices and Solid-State Circuits (EDSSC), Hsinchu, Taiwan, 18–20 October 2017; pp. 1–2. [[CrossRef](#)]
23. Chen, L.; Lain, J. Two-Dimensional Indoor Visible Light Positioning Using Smartphone Image Sensor. In Proceedings of the 2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), Taichung, Taiwan, 19–21 May 2018; pp. 1–2. [[CrossRef](#)]
24. Jeong, E.M.; Kim, D.R.; Yang, S.H.; Kim, H.S.; Son, Y.H.; Han, S.K. Estimated position error compensation in localization using visible light communication. In Proceedings of the 2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN), Da Nang, Vietnam, 2–5 July 2013; pp. 470–471. [[CrossRef](#)]
25. Xu, W.; Wang, J.; Shen, H.; Zhang, H.; You, X. Indoor Positioning for Multiphotodiode Device Using Visible-Light Communications. *IEEE Photonics J.* **2016**, *8*, 1–11. [[CrossRef](#)]
26. Zhang, X.; Duan, J.; Fu, Y.; Shi, A. Theoretical Accuracy Analysis of Indoor Visible Light Communication Positioning System Based on Received Signal Strength Indicator. *J. Lightw. Technol.* **2014**, *32*, 4180–4186. [[CrossRef](#)]
27. Li, Z.; Feng, L.; Yang, A. Fusion Based on Visible Light Positioning and Inertial Navigation Using Extended Kalman Filters. *Sensors* **2017**, *17*, 1093. [[CrossRef](#)]
28. Kim, H.; Kwon, D.; Yang, S.; Son, Y.; Han, S. Implementation of real-time indoor positioning system using carrier allocation visible light communication. In Proceedings of the 2014 12th International Conference on Optical Internet 2014 (COIN), Jeju, Korea, 27–29 August 2014; pp. 1–2. [[CrossRef](#)]
29. Kim, Y.; Hwang, J.; Lee, J.; Yoo, M. Position estimation algorithm based on tracking of received light intensity for indoor visible light communication systems. In Proceedings of the 2011 Third International Conference on Ubiquitous and Future Networks (ICUFN), Dalian, China, 15–17 June 2011; pp. 131–134. [[CrossRef](#)]
30. Zsolczai, V.; Szabó, G.; Fehér, G.; Udvary, E. Experimental investigation of multiplexing methods in Visible Light Communication system for indoor positioning. In Proceedings of the 2016 18th International Conference on Transparent Optical Networks (ICTON), Trento, Italy, 10–14 July 2016, pp. 1–4. [[CrossRef](#)]
31. Prince, G.B.; Little, T.D.C. Two-Phase Framework for Indoor Positioning Systems Using Visible Light. *Sensors* **2018**, *18*, 1917. [[CrossRef](#)] [[PubMed](#)]
32. Alam, F.; Chew, M.T.; Wenge, T.; Gupta, G.S. An Accurate Visible Light Positioning System Using Regenerated Fingerprint Database Based on Calibrated Propagation Model. *IEEE Trans. Instrum. Meas.* **2018**, 1–10. [[CrossRef](#)]
33. Konings, D.; Parr, B.; Alam, F.; Lai, E.M. Falcon: Fused Application of Light Based Positioning Coupled With Onboard Network Localization. *IEEE Access* **2018**, *6*, 36155–36167. [[CrossRef](#)]
34. Hosseinianfar, H.; Noshad, M.; Brandt-Pearce, M. Positioning for visible light communication system exploiting multipath reflections. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017; pp. 1–6. [[CrossRef](#)]
35. Gu, W.; Aminikashani, M.; Deng, P.; Kavehrad, M. Impact of Multipath Reflections on the Performance of Indoor Visible Light Positioning Systems. *J. Lightw. Technol.* **2016**, *34*, 2578–2587. [[CrossRef](#)]
36. Huang, C.; Zhang, X. Impact and feasibility of darklight LED on indoor visible light positioning system. In Proceedings of the 2017 IEEE 17th International Conference on Ubiquitous Wireless Broadband (ICUWB), Salamanca, Spain, 12–15 September 2017; pp. 1–5. [[CrossRef](#)]
37. Chen, J.; You, X. Visible light positioning and communication cooperative systems. In Proceedings of the 2017 16th International Conference on Optical Communications and Networks (ICOON), Wuzhen, China, 7–10 August 2017; pp. 1–3. [[CrossRef](#)]
38. IEEE Standard for Local and Metropolitan Area Networks—Part 15.7: Short-Range Wireless Optical Communication Using Visible Light. *IEEE Std.* **2011**, 1–309. [[CrossRef](#)]
39. Alsalami, F.M. Game theory MiniMax filter design for indoor positioning and tracking system using visible light communications. In Proceedings of the 2016 6th International Conference on Information Communication and Management (ICICM), Hatfield, UK, 29–31 October 2016; pp. 197–200. [[CrossRef](#)]

40. Yoshino, M.; Haruyama, S.; Nakagawa, M. High-accuracy positioning system using visible LED lights and image sensor. In Proceedings of the 2008 IEEE Radio and Wireless Symposium, Orlando, FL, USA, 22–24 January 2008; pp. 439–442. [[CrossRef](#)]
41. Nakazawa, Y.; Makino, H.; Nishimori, K.; Wakatsuki, D.; Komagata, H. LED-tracking and ID-estimation for indoor positioning using visible light communication. In Proceedings of the 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Busan, Korea, 27–30 October 2014; pp. 87–94. [[CrossRef](#)]
42. Tanaka, T.; Haruyama, S. New Position Detection Method Using Image Sensor and Visible Light LEDs. In Proceedings of the 2009 Second International Conference on Machine Vision, Dubai, UAE, 28–30 December 2009; pp. 150–153. [[CrossRef](#)]
43. Yang, Z.; Fang, J.; Lu, T.; Jiang, Z.L.; Chen, Z. An efficient visible light positioning method using single LED luminaire. In Proceedings of the 2017 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR), Singapore, 31 July–4 August 2017; pp. 1–2. [[CrossRef](#)]
44. Juneja, S.; Vashisth, S. Indoor positioning system using visible light communication. In Proceedings of the 2017 International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN), Gurgaon, India, 12–14 October 2017; pp. 79–83. [[CrossRef](#)]
45. Pan, W.; Hou, Y.; Xiao, S. Visible light indoor positioning based on camera with specular reflection cancellation. In Proceedings of the 2017 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR), Singapore, 31 July–4 August 2017; pp. 1–4. [[CrossRef](#)]
46. Zhang, R.; Zhong, W.; Kemao, Q. A singular value decomposition-based positioning algorithm for indoor visible light positioning system. In Proceedings of the 2017 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR), Singapore, 31 July–4 August 2017; pp. 1–2. [[CrossRef](#)]
47. Do, T.H.; Yoo, M. An in-Depth Survey of Visible Light Communication Based Positioning Systems. *Sensors* **2016**, *16*, 678. [[CrossRef](#)] [[PubMed](#)]
48. Zhuang, Y.; Hua, L.; Qi, L.; Yang, J.; Cao, P.; Cao, Y.; Wu, Y.; Thompson, J.; Haas, H. A Survey of Positioning Systems Using Visible LED Lights. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 1963–1988. [[CrossRef](#)]
49. Priyantha, N.B.; Chakraborty, A.; Balakrishnan, H. The Cricket Location-support System. In Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, Boston, MA, USA, 6–11 August 2000; ACM: New York, NY, USA, 2000; pp. 32–43. [[CrossRef](#)]
50. Danakis, C.; Afgani, M.; Povey, G.; Underwood, I.; Haas, H. Using a CMOS camera sensor for visible light communication. In Proceedings of the 2012 IEEE Globecom Workshops, Anaheim, CA, USA, 3–7 December 2012; pp. 1244–1248. [[CrossRef](#)]
51. Chavez-Burbano, P.; Guerra, V.; Rabadan, J.; Rodríguez-Esparragón, D.; Perez-Jimenez, R. Experimental Characterization of Close-Emitter Interference in an Optical Camera Communication System. *Sensors* **2017**, *17*, 1561. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).