The Utilization of a Depth Sensor in Visible light Communications

Othman Isam Younus[1], Sampath Kumar Molleti[1], Zabih Ghassemlooy[1], Joaquin Perez[2], Vicent Girbés-Juan[2], and Luis Nero Alves[3]

[1] *Optical Communications Research Group, Faculty of Engineering and Environment, Northumbria University*, Newcastle upon Tyne NE1 8ST, U.K.

[2] *Department of Electronic Engineering, Universitat de València, 46100 Burjassot, Spain*

[3] *Instituto de Telecomunicações* and *Departamento de Electrónica, Telecomunicações e Informática, Universidade de Aveiro*, 3810-193 Aveiro, Portugal.

Corresponding author: Othman Isam Younus (e-mail: othman.younus@outlook.com)

Abstract—Depth-sensing cameras are opening new applications in scientific and commercial fields. It leverages the use of infrared projected images, which can measure the depth of the scene and hence offer a three-dimensional view of the captured scene. In this paper, we investigate the performance of visible light communication with a camera-based receiver attached to a robotic arm and show that it offers a high position accuracy. The results also show that, the luminous intensity of the illuminance surface has insignificant impact on the *z* measurements within the applicable distance range of the RGB-z depth camera, and hence this may be incorporated in any VLC system designed to provide a depth estimation as well as data communication.

Keyword— Depth Sensing, RGB-z depth camera, LED, Matlab.

I. Introduction

Depth sensing devices have several advantages, which can be used in visible light communications (VLC), which is a subset of optical wireless communication (OWC) [1]. VLC uses the light-emitting diode (LED)-based lighting infrastructure for simultaneous illumination, data communications, and positioning [2], [3], [4]. In optical wireless communications, the typical optical receivers (Rxs) used are photodiodes (PDs) and image sensors (ISs), i.e., cameras. PDs can be used for a wide range of data rates (low to high) and are relatively cheap compared with IS-based Rxs. Whereas ISs with orthogonally aligned PD-array with large detection areas offer multiple Rx capabilities for detecting multiple light sources at low data rates, which are suitable mainly for the Internet of things. The standard digital PD-array (cameras) provide images with a two-dimensional (2D) pixel grid, where each pixel unit denotes a square or a dot area, which collectively form a captured frame [5]. In ISs, each pixel captures the light with an associated value based on its capturing format (usually ranging from 0-255) and has different red, blue, and green (RGB) colors.

The combination of VLC with the IS-based Rx is best known as optical camera communication (OCC), which has been used in several application areas including vehicle-to-vehicle communications, vehicle-to-infrastructure communication (i.e., traffic and street lights), device-to-device communications, etc. [4], [6]. In OCC systems, the link distance between the transmitter (Tx) and the Rx can be estimated using depth cameras. The recent developments in technology have made depth cameras cost-effective, which can be used in both commercial and industrial sectors [7].

Depth sensing devices can also be used for indoor positioning systems (IPSs). Light-based IPSs have gained significant importance over the past few years, which enhances the use of depth sensors for various applications like tracking and localizing, robotics, buildings, warehouses etc.[8]. Acquiring the depth data is still a challenging task even using machine learning given that the obtained data is not perfect because of the construction of a three-dimensional (3D) image from a 2D image [7], [9], [10]. However, there are different techniques to acquire depth information, which varies based on the level of accuracy and its cost [9].

In OCC, locating the region of interest (RoI) in the captured image is one of the critical factors.

In [11], the depth resolution in the near range using a computer-controlled positioning platform to move or rotate the depth camera with high accuracy was investigated. The depth was estimated by combining stereo and time of flight (TOF) cameras to obtain more accurate camera parameters. The scope of charge-coupled device (CCD)-based camera pairs was wrapped by the depth image taken by TOF to obtain the initial depth image. The experiment findings revealed that improved spatial resolution, higher-quality dense depth image and proper depth at edge positions were achieved compared with merely employing the TOF camera. In [12], a new technique for multi-view video generation with a 3D mesh and the view synthesis was proposed, which involved image capturing, depth wrapping and hole filling, boundary compensation, image synthesis and 3D mesh generation. The utilized RGB+Z pictures comprised of a single depth image, one or two RGB images and a PMD cube 3.0.

The authors in [13] demonstrated how a mobile robot can recognize obstacles and avoid them in an indoor environment using a stereo camera. In their algorithm, they used sum of absolute differences, in which the camera compares both images from left and right ISs and produced a disparity output as a depth image. The range of each identified object in the disparity map was estimated using a curve fitting tool, whereas Matlab® software was used for all other programming aspects and for the robot mobility.

In this paper, we investigate the impact of the input bias current of the LED mounted on a robotic arm on the depth estimation using an Intel® RealSense™ Depth Camera D435i. Using the depth images generated in Matlab® the depth of the LED is calculated by averaging pixels from one particular region.

The remainder of the paper is organized as follows: in the next two sections, the depth estimation methods and the experimental setup are presented. The numerical results and discussion are presented in Section IV. Finally, section V draws the conclusion.

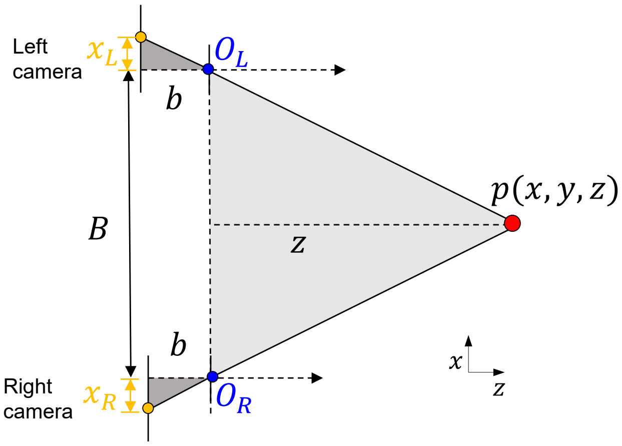


Fig. 1. The triangulation scheme with parallel cameras where a point p is projected through the optical centres O_R, O_L which yields two image points (orange) in each camera. The relative displacement of these points returns the horizontal disparity $\Delta x = x_R - x_L$. The baseline B, object distance z and image distance b affect the measured disparity.

II. System Model

a) Depth estimation using stereoscopic triangulation

Depth cameras incorporates several advanced sensing hardware to collect light information in order to generate a depth map. Different techniques have been investigated to acquire depth information. For instance, in Structured light and coded light, the system captures the 3D area using a specific pattern of light that is projected onto a scene, wherein the pattern will change in accordance with the movement of the object within the captured image field of view. A disparity matrix will be deduced to estimate the depth based on the difference between the actual and expected images [14]. It provides a lower frame rate due to the sequential projection of coded or phase-shifted patterns utilized for extracting a single depth frame [15]. A similar technique is deployed in [16], in which, the authors investigated the use of an SR300 depth camera based on a coded-light technology where triangulation between projected patterns and images captured by a dedicated sensor used to produce the depth map.

In TOF and LiDAR techniques, z is estimated by measuring the time between the emitted and reflected infrared light signal [17]. However, the method offers a limited spatial resolution due to the significant processing time required [15]. Intel® RealSense™ LiDAR Camera L515 is one of the commercial examples of this technique [17]. In the stereo depth technique, the disparity between two different imagers captured from different prespective are utilized to estimate z. The mechanism is very much like that of the human eyes, in which, each eye has its own view of vision and as the distance between eyes is fixed. Thus, generating a 3-dimensional view based on the similarities and differences between them [18]. This phenomenon is known as parallax and results in a relative displacement of respective image points from different viewpoints. These cameras can be used for both outdoor and indoor applications.

Figure 1 illustrates the stereoscopic triangulation technique, in which the sensors are coplanar. The system setup shows an object point that is projected onto both camera sensors, as represented by the orange dots in the diagram. The captured images from two different perspective with a relative distance between both camera axes is denoted as B for the baseline. The depth of the object z can be calculated using a basic mathematical triangulation, in which, the minor triangle with base of the acute scalene triangle is denoted as Δx, with the height b, whereas the major triangle has the base B and the height z.

For the estimation of the displacement, the horizontal disparity $\Delta x = x_R - x_L$, where x_R and x_L denote horizontal distances from each projected image point to the optical image centre O_R and O_L, respectively. The equality of ratios is given by:

$$\frac{z}{B} = \frac{b}{\Delta x}. \tag{1}$$

By rearrange (1), the depth information is given as:

$$z = \frac{b \times B}{\Delta x}. \tag{2}$$

Therefore, it is feasible to retrieve information about z. Likewise, if x is constant, then it is evident that by decreasing B, z reduces. Note, (*i*) for a case where the depth range long then B needs to be large; and (*ii*) (1) and (2) are only used when O_R and O_L are aligned in parallel.

b) VLC system Implementation

LED lights are mainly described based on both the luminous intesity, which represents the energy flux per solid angle, and the transmit optical power P_t. The luminous intensity associated with illuminated surfaces and normalized to the visibility. It is also used to express the brightness of an LED. However, the total energy radiated from an LED is represented by P_t. The horizontal luminouse at a point $p\ (x, y, z)$ is given by [19]:

$$E_{hr} = I_b(0) \cos^m(\emptyset) / d^2\ . \cos(\gamma), \tag{3}$$

where $I_b(0)$ is the centre luminouse intensity, the irradiance and incidence angles are denoted as $\emptyset$ and γ, respectively. Note that, the LED chip has a Lamertian pattern with an order m and is given by the semi-angle at half illuminance of an LED φ as $m = \frac{\ln 2}{\ln(\cos\varphi)}$. The distance between an LED and the detector surface is d, which given as:

$$d^2 = z^2 + h^2, \tag{4}$$

where z and h are the horizontal distance and the vertical height between the LED and the detector surface, respectively.

The received optical power received from the direct light is given as:

$$P_r = H(0). P_t\ , \tag{5}$$

where $H(0)$ is the channel DC gain within the Rx field of view and is given by:

$$H(0) = \frac{(m+1)\, A_r}{2\,\pi\, d^2} \cos^m(\emptyset)\ T_s(\gamma)\, g(\gamma) \cos(\gamma)\,, \tag{6}$$

where A_r is the area of the photodetector, $g(\gamma)$ and $T_s(\gamma)$ are the optical gain both the filter and the concentrator, respectively.

From equations (4-6), it is noted that the d and $I_b(0)$ may impact the VLC system performance. Thus, in the following sections, we aim to investigate the relationship between z

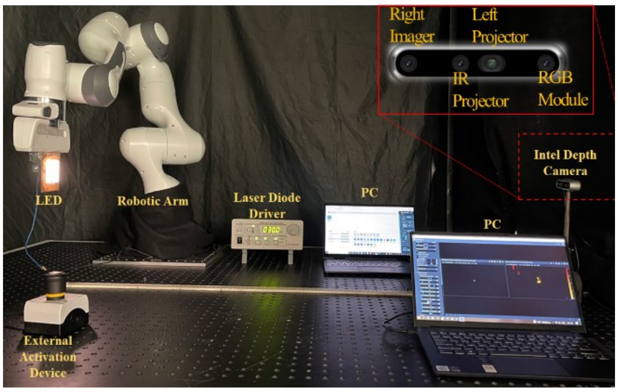


Fig. 2. A snapshot of the experimental setup of the proposed system and the depth camera used.

measurements using the depth module and the applied VLC current.

III. EXPERIMENTAL SETUP

In this work, we investigate the impact of LED drive current on the accuracy of depth measurements for a range of distances. Fig. 2 shows an indoor experimental testbed, which is used to investigate the useability of the depth camera in VLC systems under various illumination conditions.

The set-up involves a robotic arm fitted with and optical Tx (i.e., an LED) with a current driver (Newport Model 560B), and an Intel D435i depth camera used as an optical Rx. The depth camera has left and right imager, IR projector and red, green, and blue (RGB) module [20]. The IR projector is used to transmit infrared rays onto the scene prior to capturing the frames from both imagers. The depth map (i.e., the measured z at each pixel of the captured frame) is the then generated using the collected frames and the stereoscopic triangulation, see section II (a). Note that, (*i*) the Tx's position is changed by moving the computer controlled robotic arm with an accuracy of 0.1 mm [21]; (*ii*) the system is tested under a dark environment (i.e., Lux = 0); and (*iii*) the distance boundaries set are based on z from 280 mm to 2 m. The rest of key system parameters are detailed in Table II.

Table II. System parameters

Description		**Value**
Tx	LED type	Luxeon Rebel LED (SR-01-WC310)
	Tx bias current I_b	10-50 mA
Depth Rx	Rx depth module	Intel® RealSense™ D435i
	Stereo baseline	50 mm [11]
	Horizontal field of view	91.2°
	Stereo module resolution	848 × 480 pixels
	RGB camera resolution	1280 × 720 pixels
	z-accuracy (or absolute error)	≤ 2 % (up to 2 m and 80% FOV) [21]
	Minimum- z depth	Between 105 mm to 280 mm [21]
Robotic arm	Model	Franka emika robotic arm
	Power use	80 W
	Weight	18 Kg
	Precision	0.1 mm
	Maximum distance: .Sides .Top .Below ground level	 ≃ 850 mm 1190 mm 360 mm
Channel	Length	0 to 1.5 m

Algorithm I CHT for pupil boundary detection using 2-D accumulator array

Input: The raw .dng file of captured frame $F_{U\times V}$, minimum circle radius ($r_{\min}$) and maximum circle radius ($r_{\max}$)
Output: Circle radius (r_c) and centre coordinates (x_c, y_c) // comments:
for (pupil_radius = $r_{\min}$; $r \le r_{\max}$; $r = r + 1$) do
  A=zeros(rows,cols) ; // 2-D accumulator of the iris image size
  for each "Edge-point (a,b)" in edge map of $F_{U\times V}$ do
   for θ =1 ; $\theta \le 360°$; $\theta = \theta + 1$ do
   x = round ($a + r * \cos\theta$)
   y = round ($b + r * \sin\theta$)
    if (x,y) is in image bounds do
   A(x,y) = A(x,y)+1 ; // Accumulator-voting step
     end if
    end for
   end for
  Find maximum value in A:
  M=A(x',y') ; //M is maximum value in A
  Max_Array(pupil_radius) = M ;
  X_Array(pupil_radius) = x' ;
  Y_Array(pupil_radius) = y' ;
 end for
 M'= Max_Array(index) //Find maximum value
 r_c = index; x_c = X_Array(index);
 y_c = Y_Array(index) ;
 // End of CHT algorithm

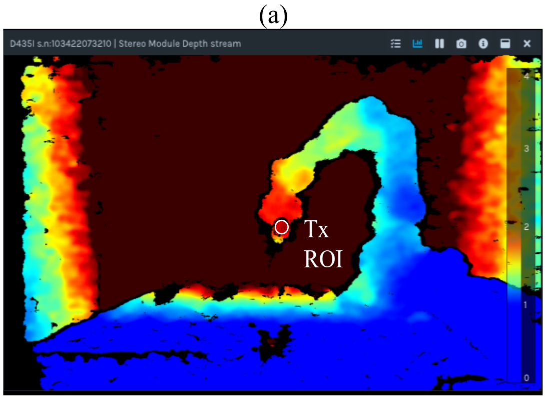
Next, the system is designed to initially detect the LED illumination footprint i.e., the region of interest (ROI). Considering that, the radiation patterns of typical a Tx LED follow Lambertian profile [22], a circular Hough transform (CHT) method is used to detect the LED illumination footprint. The extraction process is done via a voting process as described in Algorithm I. The circle description is done by applying a full sweep of θ for 360° to locate the Tx's fingerprint with a radius p and the centre coordinates of (x, y). The CHT method is implemented offline using Matlab® and by including three main steps of accumulator array computation, centre estimation, and radius estimation. Consequently, the z value is determined by averaging the captured ROI to produce the measurements.

IV. RESULTS AND DISCUSSION

The experiment is set up as proposed in section III. An example of the captured frames at I_b of 50mA and z of 1 m from both the RGB camera and it is associated depth map is depicted in Fig. 3 (a) and (b), respectively.

The impact of applied bias current (10 mA < I_b < 50 mA) is investigated for different z, and the corresponded depth and RGB frames are captured. The Tx fingerprint (i.e., ROI) is extracted using CHT as described in section III. The extracted ROI is highlight in red colour as depicted in Fig. 4. Likewise, the averaged z is measured within the Tx ROI for different I_b. Figure 5 shows the difference between the real and the estimated z for a range of I_b.

(a)



(b)

Fig. 3. The Rx captured frame of the Tx optical signal at I_b of 50mA and z of 1 m: (a) RGB camera, and (b) depth camera.

It demonstrates that increasing the distance may increase the induced error of the estimated z. For instance, a mean induced error of 512 mm is obtained at a distance of 1200mm as compared with 144 mm that is measured at 600 mm. The results also show that, the luminous intensity of the illuminance surface has insignificant impact on the z measurements within the applicable distance range of the RGB-z depth camera, and hence this may be incorporated in any VLC system designed to provide a depth estimation as well as data communication.

V. Conclusions

Depth cameras have opened new opportunities in optical-based positioning and communications. Combining the depth

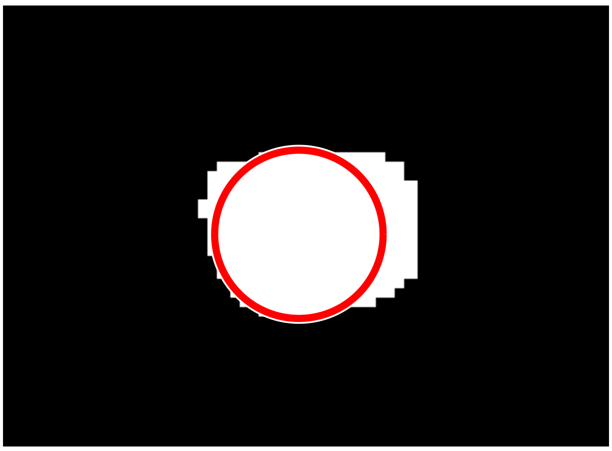


Fig. 4. The ROI extracted using the CHT algorithm. Note, that the boundaries are represented by the red colour.

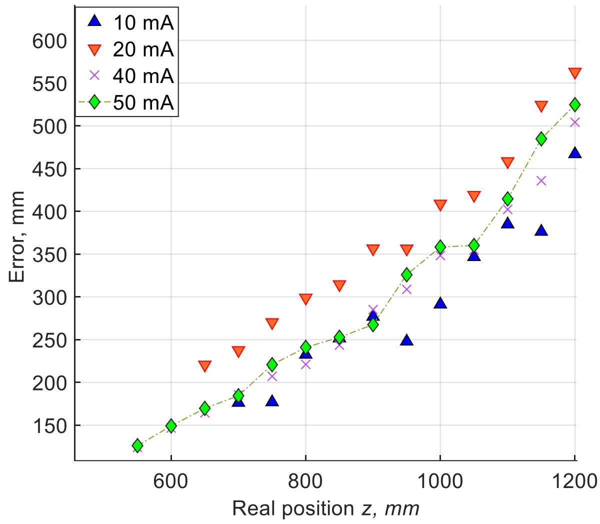


Fig. 5. The error graph between the actual distance and the depth values for a range of I_b.

information does provide a potential advantage in real-time VLC systems. In this work, we investigated the impact of the surface illuminance by changing the drive current of the Tx on the accuracy of depth measurements. The results showed that, the luminous intensity of the illuminance surface has an insignificant impact on the depth measurements within the applicable distance range of the RGB-z depth camera. Hence, the proposed scheme may be incorporated in VLC systems to provide the depth estimation as well as data communication as part of Internet of everything.

Acknowledgement

This work is supported by Northumbria University Ph.D. Scholarship, EU COST Action NEWFOCUS CA19111, and the AICAR GV/2021/074.

References

[1] X. Fernando and H. Farahneh, *Visible Light Communications*. IOP Publishing, 2019.

[2] R. Sagotra and R. Aggarwal, ‘Visible light communication’, *Int. J. Eng. Trends Technol.*, vol. 4, no. 3, pp. 403–405, 2013.

[3] E. Eso, Z. Ghassemlooy, S. Zvanovec, A. Gholami, A. Burton, N. B. Hassan, and O. I. Younus, ‘Experimental demonstration of vehicle to road side infrastructure visible light communications’, 2019, pp. 85–89.

[4] E. Eso, Z. Ghassemlooy, S. Zvanovec, J. Sathian, M. M. Abadi, and O. I. Younus, ‘Performance of Vehicular Visible Light Communications under the Effects of Atmospheric Turbulence with Aperture Averaging’, *Sensors*, vol. 21, no. 8, p. 2751, 2021.

[5] ‘Techopedia’. [Online]. Available: https://www.techopedia.com/definition/24012/pixel

[6] O. I. Younus, N. Bani Hassan, Z. Ghassemlooy, P. A. Haigh, S. Zvanovec, L. N. Alves, and H. Le Minh, ‘Data Rate Enhancement in Optical Camera Communications Using an Artificial Neural Network Equaliser’, *IEEE Access*, vol. 8, pp. 42656–42665, 2020, doi: 10.1109/ACCESS.2020.2976537.

[7] D. Pohl, S. Dorodnicov, and M. Achtelik, ‘Depth map improvements for stereo-based depth cameras on drones’, 2019, pp. 341–348.
[8] O. I. Younus, N. Chaudhary, Z. N. Chaleshtori, Z. Ghassemlooy, L. N. Alves, and S. Zvanovec, ‘The Impact of Blocking and Shadowing on the Indoor Visible Light Positioning System’, p. 6.
[9] H. Yu, G. Yang, Y. Zhou, and Z. Jin, ‘Improveing depth estimation through fusion of stereo and TOF’, 2011, pp. 563–567.
[10] R. S. Pahwa, M. N. Do, T. T. Ng, and B.-S. Hua, ‘Calibration of depth cameras using denoised depth images’, 2014, pp. 3459–3463.
[11] T.-M. Wang and Z.-C. Shih, ‘Measurement and Analysis of Depth Resolution Using Active Stereo Cameras’, *IEEE Sens. J.*, vol. 21, no. 7, pp. 9218–9230, Apr. 2021, doi: 10.1109/JSEN.2021.3054820.
[12] Hyok Song, Jisang Yoo, Sooyeong Kwak, Cheon Lee, and Byeongho Choi, ‘3D mesh and multi-view synthesis implementation using stereo cameras and a depth camera’, in *2013 3DTV Vision Beyond Depth (3DTV-CON)*, Aberdeen, United Kingdom, Oct. 2013, pp. 1–3. doi: 10.1109/3DTV.2013.6676645.
[13] Rostam Affendi Hamzah, Hasrul Nisham Rosly, and S. Hamid, ‘An obstacle detection and avoidance of a mobile robot with stereo vision camera’, in *2011 International Conference on Electronic Devices, Systems and Applications (ICEDSA)*, Kuala Lumpur, Malaysia, Apr. 2011, pp. 104–108. doi: 10.1109/ICEDSA.2011.5959032.
[14] ‘Robotics Tomorrow’. [Online]. Available: https://www.roboticstomorrow.com/article/2018/04/what-is-structured-light-imaging/11821
[15] A. Kadambi, A. Bhandari, and R. Raskar, ‘3D Depth Cameras in Vision: Benefits and Limitations of the Hardware’, in *Computer Vision and Machine Learning with RGB-D Sensors*, L. Shao, J. Han, P. Kohli, and Z. Zhang, Eds. Cham: Springer International Publishing, 2014, pp. 3–26. doi: 10.1007/978-3-319-08651-4_1.
[16] A. Zabatani, V. Surazhsky, E. Sperling, S. B. Moshe, O. Menashe, D. H. Silver, Z. Karni, *et al.*, ‘Intel® RealSense™ SR300 Coded Light Depth Camera’, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 10, pp. 2333–2345, Oct. 2020, doi: 10.1109/TPAMI.2019.2915841.
[17] ‘Intel® RealSense™ Depth Camera D435i’. [Online]. Available: 1. https://www.intelrealsense.com/beginners-guide-to-depth/
[18] R. A. Hamzah, H. N. Rosly, and S. Hamid, ‘An obstacle detection and avoidance of a mobile robot with stereo vision camera’, 2011, pp. 104–108.
[19] T. Komine and M. Nakagawa, ‘Fundamental analysis for visible-light communication system using LED lights’, *IEEE Trans. Consum. Electron.*, vol. 50, no. 1, pp. 100–107, Feb. 2004, doi: 10.1109/TCE.2004.1277847.
[20] Y. Benkhoui, T. El Korchi, and L. Reinhold, ‘UAS-Based Crack Detection Using Stereo Cameras: A Comparative Study’, 2019, pp. 1031–1035.
[21] ‘Intel® RealSense™ Product Family D400 series Datasheet’. Jan. 2020. [Online]. Available: https://www.intelrealsense.com/wp-content/uploads/2020/06/Intel-RealSense-D400-Series-Datasheet-June-2020.pdf
[22] I. Moreno and C.-C. Sun, ‘Modeling the radiation pattern of LEDs’, *Opt. Express*, vol. 16, no. 3, pp. 1808–1819, Feb. 2008, doi: 10.1364/OE.16.001808.