GIPS: Geometry-Inspired Passive ToF Sensing for 3D Depth Reconstruction

Faisal Ahmed[†], Miguel Heredia Conde[†‡], and Paula López Martínez[‡]
[†]Center for Sensorsystems (ZESS), University of Siegen, Paul-Bonatz-Straße 9-11, 57076 Siegen, Germany
[‡] CiTIUS, University of Santiago de Compostela 15782, Santiago, Spain
Emails: faisal.ahmed@uni-siegen.de • heredia@zess.uni-siegen.de • p.lopez@usc.es

***Abstract*—The motivation for this work lies in the ubiquitous lighting infrastructure that surrounds us, which has been repurposed as a medium for optical wireless communication (OWC). This technology has enabled the development of passive 3D imaging. However, the fundamental problem with passive Time-of-Flight (ToF) imaging is the unknown source location, which generates a *chicken and egg* problem that renders passive ToF unusable in practice. In this work, we propose computational methods to solve this problem. We present a novel algorithm with a gradient descent approach for jointly estimating the source location and retrieving the correct depth information of the scene. In each iteration, the bistatic configuration is used as a basic framework while seeking local planarity to constrain the source location. The performance of our algorithm is evaluated in terms of source location estimation error and depth reconstruction error for two usual plane orientations through numerical simulations. Simulation results confirm the ability of the method to jointly retrieve the scene depth and source location. This work has huge potential for next-generation wireless networks, e. g., 6G, and paves the way towards 3D reconstruction in multiple practical applications.**

***Index Terms*—Bistatic passive sensing, gradient descent, depth estimation, and blind source localization.**

I. Introduction

3D source localization in passive imaging is a challenging task. Prior work relies on the assumption that the source location is known, which does not hold in general [1], [2]. In this context, source localization can be viewed as a *chicken and egg* problem, where the position of the source cannot be determined unless the geometry of the scene is known and the depth image cannot be retrieved correctly without the 3D source position. Blind source localization refers to the problem of determining the position of the illumination source without any prior knowledge [3]. To this end, we propose a novel algorithm that leverages the sensing potential of a bistatic setup and geometric priors to effortlessly determine the position of a 3D source. This work paves the way for passive Time-of-Flight (ToF) imaging to be operational in real applications requiring 3D depth reconstruction.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No: 860370 (MENELAOS[NT]). This work has received funding from the Spanish Ministry of Science, Innovation and Universities under grant PID2021-128009OB-C32 and from Xunta de Galicia-Consellería de Cultura, Educación e Universidade Accreditations 2019-2022 ED431G-2019/04 and 2021-2024 ED431C2021/048 (ERDF/FEDER programme).

Recently, interest in passive 3D imaging has been stimulated by our previous work [1], [2], which exploited existing optical wireless communication (OWC) infrastructure. The quest for 3D depth reconstruction has become a prevalent topic of research in the computer vision and signal processing communities, with potential applications in autonomous cars, robots, and indoor sensing. Our work aligns well with the global trend of "re-purposing optical paradigms" that has engulfed several scientific disciplines (e. g., indoor sensing) in recent decades [4]. We exploited an array of *photonic mixer device* (PMD) correlating pixels in this work, which perform the mixing between the received optical signal and the reference one [5].

State-of-the-art ToF-based 3D cameras use co-located illumination sources to extract depth by exploiting time or phase delay information. Differently, the illustration in Fig. 1a presents an alternative approach, known as a bistatic configuration that exploits existing OWC infrastructure and yields a passive modality. Fig. 1b demonstrates the resulting 3D ellipsoidal model where $P_{\mathrm{E}}, P_{\mathrm{T}}$, and P_{R} are the 3D locations of the emitter, target, and receiver, respectively. The ellipsoid has two foci, such that the sum of the distances from any point on the surface to the two foci is constant. For each pixel, the solution set lies on this 3D ellipsoid as, demonstrated in Fig. 1b. Our previous work [1] requires knowledge of the emitter location P_{E} to estimate the correct distance to target, which is not generally available. In this work, we show how geometrical cues can be used to estimate P_{E} from the measurements.

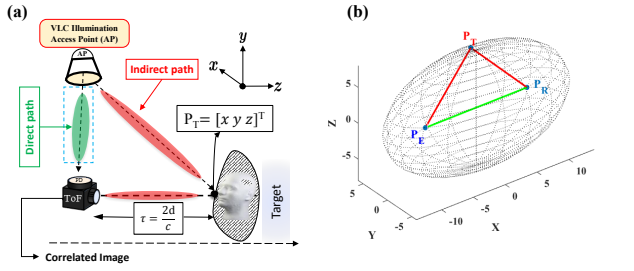

Fig. 1. (a) Bistatic Passive-ToF imaging configuration. (b) For each pixel, the solution set lies on a 3D ellipsoid.

In 3D reconstruction, scene priors are valuable in addition to geometric cues. The human visual system makes unconscious assumptions about the 3D object based on common observations, such as its simplicity, symmetry, and planarity [6]–[8]. To streamline the convex optimization problem, we exploit

scene priors. More specifically, we demonstrate the method leveraging local horizontal and vertical planarity, because this is the most common case in practice, referred to as the *Manhattan world*. Prior knowledge of the scene, particularly the *Manhattan world* assumption, has been demonstrated to enhance 3D reconstruction tasks as reported in [9]. In practice, any complex structure can be approximated reasonably well by using small planar components. Indoor scenes (walls and floors) [10] and self-driving cars (road surfaces) are two examples of such applications.

Key Contributions: This work demonstrates an approach for solving the blind source localization in bistatic passive ToF imaging that follows the principles of [1].

1) We present a novel *passive framework* that combines geometric prior and optimization methods for jointly estimating the target and source position.
2) We report how blind source location plays a key role in the passive modality for accurate depth reconstruction.
3) We propose a novel bistatic passive gradient descent (BP-GD) algorithm that uses an optimization approach to reconstruct a 3D image in a passive ToF setting with an *unknown source location*.

II. Related Work

3D Active Imaging: Conventional depth imaging methods, such as structured light [11] and ToF imaging [12], are active sensing techniques that acquire depth information by capturing an image using a camera and an integrated illumination source. While these methods have made significant progress in recent times, however, they still face key challenges. Structured-light sources cannot be used in some applications, e. g., autonomous drones, owing to cost and power constraints. Despite the obvious advantages, active ToF imaging has its own set of problems, such as power-hungry sources, background disturbance, and temperature-dependent measurements. Unlike recent approaches, we propose a passive modality that is suitable for power-constrained applications that exploit opportunity illuminators, e. g., visible light communication (VLC) sources, for distilling depth information. The latter is embedded in reflections in traditional active ToF imaging methods, where the illumination source is integrated within the sensor. It has been argued in [1], [2] that passive ToF imaging requires a receiver-only sensor in order to obtain depth reconstruction.

3D Passive Imaging: Many approaches have recently leveraged photon delays to infer depth information in passive imaging. The photon bunching phenomena in thermal light can be leveraged to retrieve depth, as discussed in [13], which builds on the earlier works of Hanbury Brown and Twiss [14]. This approach has certain downsides, such as the need for optical conditioning in the sensing pipeline and the cost and bulkiness of the system itself. Photometric stereo (PS) imaging is a prominent example of passive imaging [15]. PS imaging, pioneered by Woodham in 1980 [16], uses a fixed camera perspective and multiple light directions to capture 3D images of an object. Conventional PS approaches assume light directions and intensities are constant, known or unknown among all frames. PS imaging takes the source position fixed with respect to the scene. However, there is another catch: this technique can only be used in a completely dark environment. The artifacts in the surface reconstruction problem are caused mainly by the discontinuity or the sudden shift in intensity between objects and depth leading to a zero value. [17] tackles the problem of the hidden source position. This framework uses aperture masking interferometry for passive imaging, which is limited by signal-to-noise ratio (SNR) considerations, and exploits costly ultrafast detectors, as PS does not require ultra detectors [15].

None of the existing approaches solves the entangled problem of simultaneously reconstructing the scene geometry and the source location in passive 3D imaging. Even supposing a known initial position of the source, in dynamic systems, this could change over time. Thus, relying on a known source location is not a realistic assumption in general. Our method makes use of total round-trip time, thus the depth, which is measured in active ToF methods, here, is termed as *incorrect depth* due to the unknown source location. We develop the necessary mathematical machinery to ensure that the novel passive imaging can determine the source position. Our proposed algorithm combines a bistatic formulation and an iterative optimization strategy that seeks local planarity to retrieve the unknown source location and correct the depth.

III. Proposed Methodology

The proposed bistatic geometry is demonstrated in Fig. 1a. The LED source illuminates the scene with the modulated light. Moreover, a lens forms an image on the camera's focal plane. For generating the ground truth data for our numerical experiment, a simple model of the local target geometry is first defined. We generated 3D plane points, $\{P_i\}_{i=1}^{N}$, by exploiting the camera lens normals for the N pixels. The camera is located at a known position and, to generate measurements synthetically, we defined a ground-truth (GT) emitter position and calculated the incorrect depth from the camera measurements. This depth and an initial guess of the emitter location are inputs to our proposed algorithm for determining the source position and correcting the depth.

A. General Least-Squares for Plane Fitting

In this work, we have limited our attention to a planar region whose points follow equation (1). We adopt Hesse's normal form for modeling a planar patch of 3D points. A plane can be defined as:

$$Z = \alpha X + \beta Y + \gamma \tag{1}$$

where the normal vector is represented by $\vec{x} = (\alpha, \beta, \gamma)$ and $Z = [z_1, z_2, \ldots, z_N]^T$, $X = [x_1, x_2, \ldots, x_N]^T$, and $Y = [y_1, y_2, \ldots, y_N]^T$ are the 3D plane points. Here, N represents the number of plane points given by the number of pixels in the array. Given a collection of points, (1) is translated as a linear system of equations $Z = \boldsymbol{A}\vec{x}$. The matrix $\boldsymbol{A}$ is constructed as $\boldsymbol{A} = [X, Y, O]$, where $O = [1, 1, \ldots, 1]^T$, and $\boldsymbol{A} \in \mathbb{R}^{N\times 3}$. As a result, the plane fitting problem becomes an overdetermined system, as $N \gg 3$, and the normal

vector $\vec{x}$ can be estimated from Z by the Moore-Penrose pseudoinverse $\boldsymbol{A}^{\dagger} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$. Here $\hat{\vec{x}}$ is the least squares (LS) solution of the system and we obtain a fitted plane model by substituting these values in (1).

B. Ellipsoid Localization Geometry

To determine the bistatic range, the flight time of the optical signal emitted from an unknown source is multiplied by the speed of light. This gives the sum of the distances from the emitter to the target, and then to the ToF sensor. Additionally, it is important to subtract the baseline distance from the source to the receiver. This induces a 3D ellipsoid (see Fig.1b), where a target satisfying the resulting bistatic range may be located in three-dimensional space. The emitter and ToF sensor locations are its foci. We obtained the observation direction vectors of each pixel, thanks to the lens calibration. For known foci, the intersection of the ellipsoid and the direction vectors yield the 3D target points and, thus, the correct depth. The 3D target location can be formulated as follows:

$$P_{\mathrm{T}} = P_{\mathrm{R}} + \vec{n}_{\mathrm{R}}\mathrm{d}_{\mathrm{RT}} \tag{2}$$

where P_{T} and P_{R} define the 3D target and receiver points and $\vec{n}_{\mathrm{R}}$ is the unit vector along the observation direction, and d_{RT} is the distance between the receiver and the target. The 3D ellipsoid is defined by,

$$\mathrm{d} = \mathrm{d}_{\mathrm{ET}} + \mathrm{d}_{\mathrm{RT}} - \mathrm{d}_{\mathrm{ER}} \tag{3}$$

where d defines the total distance, here named as *incorrect depth*. $\mathrm{d}_{\mathrm{ET}} = \|P_{\mathrm{E}} - P_{\mathrm{T}}\|_2$ is the Euclidean distance between the emitter and the target. $\mathrm{d}_{\mathrm{ER}} = \|P_{\mathrm{E}} - P_{\mathrm{R}}\|_2$ is the baseline distance. Equation (3) is rearranged and we break down (2) into 3D coordinate components and plug it into (3) for further derivation. For brevity, the mathematical derivation has been omitted. A previous version of the derivations, supposing known source location and eliminating the effect of d_{ER} by calibration, can be found in [1]. In this work, we consider the effect of baseline distance, d_{ER}. The corrected depth, $\mathrm{d}^{'} = f_{\mathrm{d}_{\mathrm{RT}}}(\mathrm{d}, P_{\mathrm{E}})$, is a function of the incorrect depth d and P_{E} is the 3D position of an emitter. G is the scalar product of $\langle(P_{\mathrm{E}} - P_{\mathrm{R}}), \vec{n}_{\mathrm{R}}\rangle$, as follows,

$$f_{\mathrm{d}_{\mathrm{RT}}}(\mathrm{d}, P_{\mathrm{E}}) = \frac{\mathrm{d}^2 + 2\mathrm{d}\mathrm{d}_{\mathrm{ER}}}{2\mathrm{d} + 2\mathrm{d}_{\mathrm{ER}} - 2\mathrm{G}} \tag{4}$$

C. Proposed Algorithm

Let $\mathcal{C}(f_{\mathrm{d}_{\mathrm{RT}}}(\mathrm{d}, P_{\mathrm{E}}), d_{\mathrm{Fit}})$ denotes a cost function that is continuous on P_{E}. Where d_{Fit} is the distance to each point in the fitted plane. The cost function can then be minimized by applying gradient descent with respect to the parameter, P_{E}. The number of iterations to reach a local minimum is dependent on the step size η. Namely, we go down over the surface formed by the cost function until we reach the bottom of a valley. Typically, the mean squared error (MSE) is used as the cost function. This calculates how far off a model is from the ground truth. Note that we do not have access to the GT values. Instead, we obtain them by applying plane fitting to the corrected depth value at each iteration. The sum of the squared differences between the corrected depth given by $f_{\mathrm{d}_{\mathrm{RT}}}$ and the fitted plane values d_{Fit} is the chosen cost function, as given below,

$$\mathcal{C}(f_{\mathrm{d}_{\mathrm{RT}}}(\mathrm{d}, P_{\mathrm{E}}), d_{\mathrm{Fit}}) = \frac{1}{N}\sum_{n=1}^{N}(f_{\mathrm{d}_{\mathrm{RT}}}(\mathrm{d}, P_{\mathrm{E}}) - d_{\mathrm{Fit}})^2 \tag{5}$$

We have proposed an iterative extension of our bistatic algorithm by introducing an additional step of gradient descent to estimate the emitter position and correct the depth. To compute the gradient of a cost function *w.r.t* the 3D coordinate components of the emitter location, P_{E}, and estimate how the function changes with each component independently while keeping the other constant. The idea of the BP-GD Algorithm 1 is an iterative approach for solving the ellipsoid localization problem with an unknown source location in passive imaging. The algorithm involves iterating the following two steps:

1) First, the corrected depth is computed using the bistatic algorithm via (4). Next, this corrected depth is used to fit a plane using (1), where the coefficients are unknown. We estimate the coefficients using LS minimization. The cost function is affected by the plane fitting process and is highly dependent on the emitter location.
2) We compute the gradient of the cost function and use it to update the emitter location, starting from a given initial guess. Each location update guarantees either a reduction in MSE or a minor change, ensuring a monotonic reduction in MSE and convergence to a local minimum.

For our validation study, we rely on realistic simulations of 3D target geometry points in unknown source position settings. To achieve this, we use the GT emitter location and introduce noise with different variances to model the uncertainties in the initial guess, which may come from a probabilistic filter. We investigate various variances to derive statistical performance indicators that account for varying degrees of uncertainty in the initial guess, thereby enhancing the reliability of our analysis.

Algorithm 1 Blind 3D Source Localization using Bistatic Passive Algorithm with Gradient Descent

Inputs: η, $P_{\mathrm{E}}^{(0)}$, P_{R}, $\vec{n}_{\mathrm{R}}$, d
Result: $P_{\mathrm{E}}^{(K)}$, $\mathcal{C}(P_{\mathrm{E}}^{(K)})$; $d_{\mathrm{RT}}^{(K)}$
for $k \leftarrow 1$ to K **do**
  Compute corrected depth: $d_{\mathrm{RT}}^{(k)}$ using (4)
  Construct matrix $\mathbf{A}$
  Calculate estimate: $\hat{\vec{x}}^{k} = \mathbf{A}^{\dagger k}\vec{y}^{k}$
  Updated plane model: plane fitting using (1)
  Compute the cost function: $\mathcal{C}(P_{\mathrm{E}}^{(k-1)}) \leftarrow \frac{1}{N}\sum_{n=1}^{N}(f_{\mathrm{d}_{\mathrm{RT}}}(\mathrm{d}, P_{\mathrm{E}}^{(k-1)}) - d_{\mathrm{Fit}}^{(k)})^2$
  Estimate the gradient of the cost function: $\frac{\partial\mathcal{C}(P_{\mathrm{E}})}{\partial P_{\mathrm{E}}}$
  Update the emitter location: $P_{\mathrm{E}}^{(k)} \leftarrow P_{\mathrm{E}}^{(k-1)} - \eta\frac{\partial\mathcal{C}(P_{\mathrm{E}})}{\partial P_{\mathrm{E}}}$
end for

IV. Results and Discussion

A number of numerical simulations are carried out to validate our novel algorithm over synthetically generated passive

ToF data. We consider two target cases, namely vertical and horizontal planes. We define the emitter, receiver, and target points to simulate the fully passive setup. In this case, the camera lens normals are used to attain the target 3D plane points and we leverage the bistatic geometry (3) to get the incorrect depth, emulating real sensor measurements. The receiver is located at the origin of the coordinate system P_{R} =(0,0,0) cm and the emitter is placed at P_{E} =(2,30,20) cm as GT location for both (vertical and horizontal) planes, respectively. Firstly, the GT depth is generated using the l_2 distance between the target and the receiver. Fig. 2 shows the GT depth, incorrect depth using (4), and the depth error w.r.t GT for the vertical plane. The maximum depth error is of 14.3 cm. For both vertical and horizontal plane cases we use the same approach to obtain the initial guess. The initial guess is drawn from the normal distribution $\mathcal{N}(P_{\mathrm{E}}, \sigma^2)$ with $\sigma^2 \in \{5, 10, 20\}$. We carried out 12 realizations to obtain statistical metrics. For the worst-case scenario ($\sigma^2 = 20$), we report representative cases of best, medium, and worst performance based on the emitter position error. Fig. 3 shows the incorrect depth error with respect to the GT and the depth reconstruction error attained by the proposed approach, which is in the sub-centimetric range for the best case. The depth reconstruction absolute error is of 1.3 cm for the best case, 1.55 cm for the medium, and 3 cm for the worst. Fig. 4 shows the cost function value and how the estimated emitter location errors decrease with the number of iterations. In the best case, the largest depth reconstruction error values are around 1.3 cm as witnessed by the first row of plots of Fig. 3. The indirect depth deviates as the emitter location changes.

Fig. 5 demonstrates the GT depth and incorrect depth for the horizontal plane. The image size is limited by half of the pixel array, 100×640 pixels due to the fact that only the lower half of the pixel array observes the horizontal plane. We restricted 100 pixels due to infinite distance. The horizontal plane is located -9 cm below the camera. The incorrect depth is obtained by applying (3) to the depth reconstruction obtained from (4). We present the depth reconstruction for the horizontal plane in Fig. 6. The depth reconstruction absolute error is of 1.2 cm for the best case, 1.8 cm for the medium, and 2.3 cm for the worst case. Fig. 7 depicts the emitter estimation error and the cost function versus the number of iterations for the horizontal plane. It can be seen that the source error is decreased 4.2 cm and that the cost function follows the same monotonically-decreasing pattern as the emitter position error. Our indoor infrastructure consists largely of a number of planes, our method is able to reconstruct the target depth via the proposed passive algorithm. To evaluate its robustness, we added Gaussian noise to the measurements and also controlled the baseline distance, $\mathrm{d_{ER}}$, for 0 dB SNR value. The depth error is evaluated in terms of root-mean-square error (RMSE) and we attained -13 dB and -11.2 dB for the best case of the vertical plane and horizontal plane, respectively, as shown in Fig. 8. The vertical plane case exhibits a lower depth reconstruction error compared to the horizontal plane, likely due to the directional effect of the emitter on the latter. As the number of iterations increases, the cost function shows a linear decay, until an accurate reconstruction is attained.

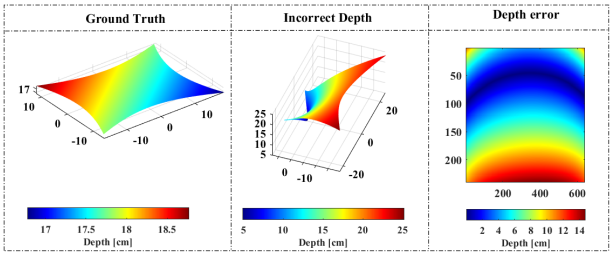

Fig. 2. 3D imaging from simulated passive ToF measurements of the vertical plane. Ground truth depth and absolute depth error with respect to GT.

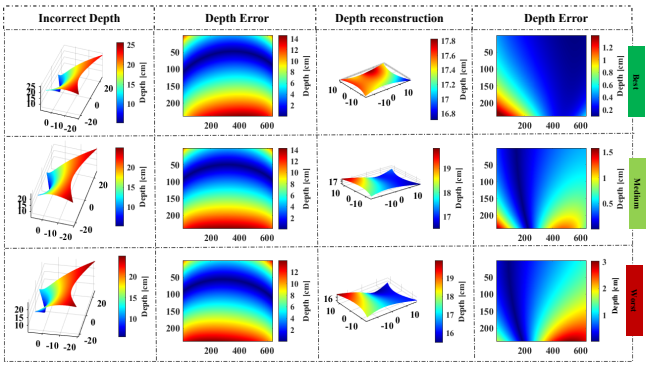

Fig. 3. Qualitative results of our approach on the vertical plane. The incorrect depth is based on the estimated emitter position (column 1). The corresponding absolute depth error is given in column 2. The depth reconstruction (column 3) and the corresponding absolute depth error plot (column 4) are shown with $\sigma^2 = 20$ at best, medium, and worst cases.

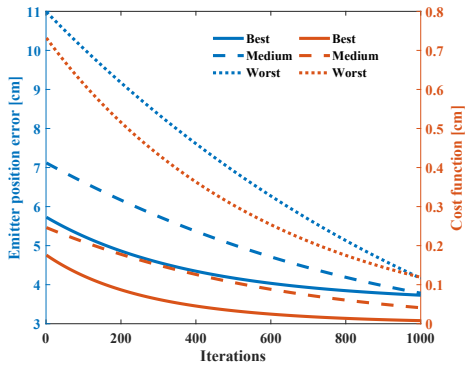

Fig. 4. Vertical plane emitter location error and cost function values versus iteration number at $\sigma^2 = 20$ for best, medium, and worst cases.

V. Conclusions and Future Work

In this work, we address the 3D ellipsoid-restricted localization problem that arises in passive ToF imaging when relying on a bistatic setup and when the emitter location is unknown. The proposed method exploits local planarity, present in most real scenarios. Our method is designed to estimate the 3D geometry of the scene and the emitter location simultaneously. The obtained results complement existing efforts to develop passive imaging technologies. Specifically, we used iterative

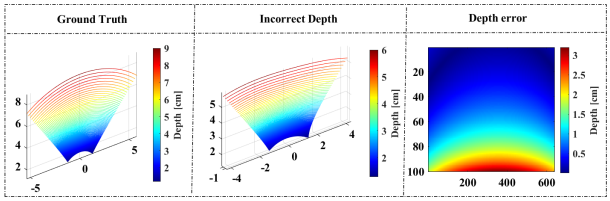

Fig. 5. Ground truth depth and absolute depth error for the horizontal plane with respect to GT.

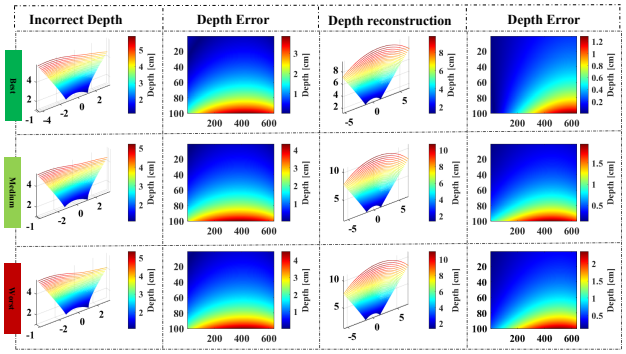

Fig. 6. Qualitative results of our approach on the horizontal plane. The incorrect depth is based on the estimated emitter position (column 1). The corresponding depth error is given in column 2. The depth reconstruction (column 3) and the corresponding error map (column 4) are shown with $\sigma^2 = 20$ at best, medium, and worst cases.

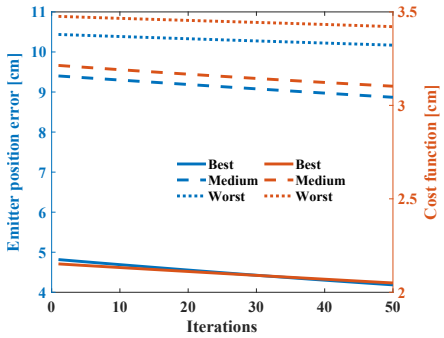

Fig. 7. Horizontal plane emitter location error and cost function values versus iteration number at $\sigma^2 = 20$ for best, medium, and worst cases.

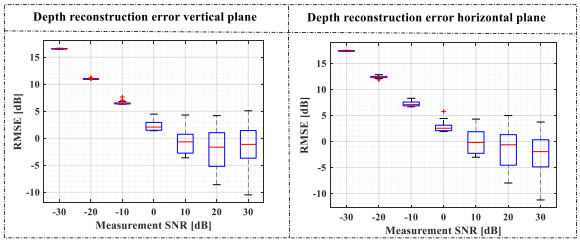

Fig. 8. Depth reconstruction error for the vertical and horizontal planes with different measurement SNR for the best case.

optimization to predict the location of the source. To achieve this, we combined a step of correction of the depth estimates given the current estimate of the bistatic geometry with another one updating the estimate of the source via gradient descent. This approach enabled us to accurately recover the source location and to correct the depth reconstruction of the scene. Experiments with synthetic data conducted in two plane configurations showed that depth reconstruction errors below 2 cm were achieved in 96% of the cases and the source location could be estimated with 3.5 cm position error. Our algorithm is used for synthetic plane scenes. Future work includes validating the algorithm using experimental data from a bistatic passive ToF camera and the development of a method for locating planar regions in the scene.

REFERENCES

[1] F. Ahmed, M. Heredia Conde, P. López Martínez, T. Kerstein, and B. Buxbaum, "Pseudo-passive time-of-flight imaging: Simultaneous illumination, communication, and 3D sensing," *IEEE Sensors Journal*, vol. 22, no. 21, pp. 21218–21231, 2022.

[2] F. Ahmed, M. Heredia Conde, and O. Loffeld, "Pseudo-passive indoor ToF sensing exploiting visible light communication sources," in *2021 IEEE Sensors*, pp. 1–4, IEEE, 2021.

[3] I. B. F. De Almeida, M. Chafii, A. Nimr, and G. Fettweis, "Blind transmitter localization in wireless sensor networks: A deep learning approach," in *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1241–1247, IEEE, 2021.

[4] H. Haas, J. Elmirghani, and I. White, "Optical wireless communication," 2020.

[5] R. Schwarte, "Method and apparatus for determining the phase and/or amplitude information of an electromagnetic wave for photomixing," May 30 2006. US Patent 7,053,357.

[6] Y. Li, Z. Pizlo, and R. M. Steinman, "A computational model that recovers the 3D shape of an object from a single 2D retinal representation," *Vision research*, vol. 49, no. 9, pp. 979–991, 2009.

[7] Z. Pizlo and A. K. Stevenson, "Shape constancy from novel views," *Perception & Psychophysics*, vol. 61, pp. 1299–1307, 1999.

[8] Z. Pizlo, "Perception viewed as an inverse problem," *Vision research*, vol. 41, no. 24, pp. 3145–3161, 2001.

[9] E. Delage, H. Lee, and A. Y. Ng, "Automatic single-image 3D reconstructions of indoor Manhattan world scenes," in *Robotics Research: Results of the 12th International Symposium ISRR*, pp. 305–321, Springer, 2007.

[10] K. Huang, Y. Wang, Z. Zhou, T. Ding, S. Gao, and Y. Ma, "Learning to parse wireframes in images of man-made environments," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 626–635, 2018.

[11] J. Geng, "Structured-light 3D surface imaging: a tutorial," *Advances in Optics and Photonics*, vol. 3, no. 2, pp. 128–160, 2011.

[12] A. Bhandari, M. Heredia Conde, and O. Loffeld, "One-bit time-resolved imaging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 7, pp. 1630–1641, 2020.

[13] F. Wagner, F. Schiffers, F. Willomitzer, O. Cossairt, and A. Velten, "Intensity interferometry-based 3D imaging," *Optics Express*, vol. 29, no. 4, pp. 4733–4745, 2021.

[14] R. Brown and R. Q. Twiss, "Correlation between photons in two coherent beams of light," *Nature*, vol. 177, no. 4497, pp. 27–29, 1956.

[15] E. L. Francois, J. Herrnsdorf, J. J. D. McKendry, L. Broadbent, G. Wright, M. D. Dawson, and M. J. Strain, "Synchronization-free top-down illumination photometric stereo imaging using light-emitting diodes and a mobile device," *Opt. Express*, vol. 29, pp. 1502–1515, Jan 2021.

[16] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical engineering*, vol. 19, no. 1, pp. 139–144, 1980.

[17] J. Boger-Lombard and O. Katz, "Passive optical time-of-flight for non line-of-sight localization," *Nature communications*, vol. 10, no. 1, pp. 1–9, 2019.