

An Easy Method for Extrinsic Calibration of Camera and Time-of-Flight Sensor

Tianyou Zhang¹, Jing Liu¹, Dragos Axinte¹ and Xin Dong¹

Abstract—A multi-zone (typically 8×8) time-of-flight (ToF) sensor offers a low-cost, low-power, and compact solution for range measurement, making it ideal for specialized robotic applications. However, its low resolution limits its usability. Pairing a ToF sensor with a camera enhances depth perception and can solve the unscaled metric problem in mono depth estimation. Advances in deep learning further enable high-quality depth map reconstruction from ToF-camera data, providing a cost-effective alternative. However, accurate ToF-camera calibration remains a challenge due to ToF sensor’s coarse depth output.

This work presents a simple yet effective method for the extrinsic calibration of a ToF sensor with an RGB camera using only a chessboard and two whiteboards. A tailored two-plane fitting algorithm is proposed specifically for the ToF sensor. Moreover, our approach leverages parallel lines with vanishing points and geometric constraints from plane intersections. This eliminates the need for robotic arm movements or SLAM-based sensor pose reconstruction, significantly reducing complexity while maintaining high accuracy. Experimental results demonstrate that our method lowers the root mean square (RMS) depth difference from 96.59 mm to 67.89 mm, underscoring its effectiveness in practical applications. Code is publicly available in <https://github.com/Tianyou-Nottingham/ToF-Camera-Calibration>.

I. INTRODUCTION

Depth perception plays a crucial role in robotic applications, particularly in Simultaneous Localization and Mapping (SLAM)[1], human-robot interaction[2], and autonomous navigation[3]. Among the commonly used range sensing technologies, LiDAR provides high-resolution depth perception through laser-based measurement. However, its high cost and large size make it impractical for special robotic applications. Other alternative depth sensing technologies can address these limitations. Stereo vision systems offer a cost-effective and lightweight solution that is relatively easy to calibrate, but their performance is highly dependent on ambient illumination conditions. Structured light sensors provide improved depth accuracy but remain bulky and expensive, making them less suitable for applications such as continuum robots. In contrast, Time-of-Flight (ToF) sensors share the fundamental working principle of LiDAR while offering significantly lower energy consumption, reduced cost, and a more compact form factor. These advantages have led to their widespread use in applications such as mobile device autofocus and obstacle avoidance.

Despite these benefits, ToF sensors suffer from a significant drawback: low spatial resolution. Most commercial

¹The authors are with Faculty of Mechanical, Materials and Manufacturing Engineering, University of Nottingham, NG8 1BB, The United Kingdom. [tianyou.zhang|eexj154|dragos.axinte|xin.dong]@nottingham.ac.uk

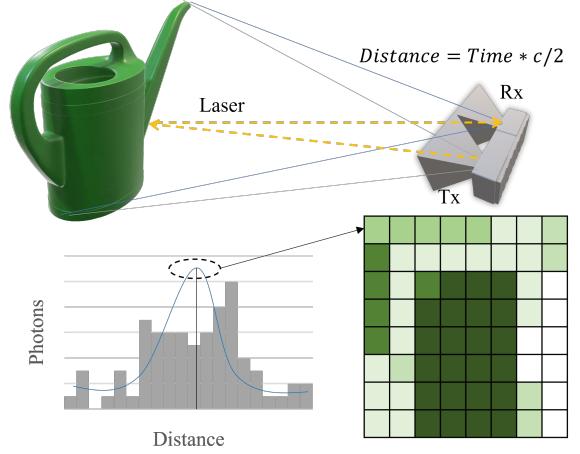


Fig. 1: Simple explanation for ToF sensor: zone of the receiver counts the photons number from emitter reflected by object’s surface (*Top*), fitting a curve or taking the mode as the value for the zone (*Bottom*).

ToF sensors offer a resolution of 8×8 or lower, which limits their ability to provide fine-grained depth information. Furthermore, ToF sensors measure the expected value of a depth distribution rather than precise point-wise depth values, meaning that they capture only an approximate depth for a given region rather than the exact depth of individual points. Consequently, ToF sensors alone are insufficient for high-precision depth applications such as SLAM or pose estimation, particularly when used in cooperation with an RGB camera.

Recent studies have employed ToF sensor for planar surface detection in robotic applications, including object placement[4] and planarity deviation analysis[5]. Furthermore, ToF sensors have been widely utilized in high-level tasks such as depth estimation[6], 3D surface reconstruction[7], and SLAM[8] with the fusion of camera. However, extrinsic calibration remains a prerequisite for ensuring the success of these tasks.

Some works [6] leverage deep learning-based methods to achieve depth estimation performance by ToF-camera comparable to LiDAR-camera systems. By training supervised models on high-quality depth maps (e.g., those obtained from Intel RealSense sensors), lightweight ToF-camera pairs can be utilized for advanced robotic tasks[8]. However, to ensure accurate data collection and model inference, precise

extrinsic calibration between the ToF sensor and the RGB camera is essential. Due to the low resolution and spatial averaging characteristics of ToF sensors, traditional calibration approaches often rely on additional external mechanisms, such as RGB-D SLAM[6] or robotic arm trajectory tracking[9].

In this study, we propose a novel extrinsic calibration method for a ToF sensor (ST VL53L8A¹, 8 × 8 resolution) and an RGB camera (Intel RealSense D435i²) that leverages geometric constraints without requiring external tracking or SLAM-based reconstruction. Our method requires only a single calibration target composed of three mutually perpendicular planes, which can also be replaced by the common building corner. By capturing a series of RGB images and corresponding ToF depth data, we achieve 6-degree-of-freedom (6-DoF) extrinsic calibration in a straightforward and efficient manner.

To accomplish this, we introduce a 2-plane fitting algorithm tailored to sparse ToF depth data. While conventional ToF plane-fitting methods typically estimate a single plane[4], our approach extends this to simultaneously fit two planes, despite the limited number of depth points, thereby enhancing calibration efficiency.

The key contributions of this work are: 1) A simple yet efficient extrinsic calibration method for ToF-RGB camera pairs that requires only a single calibration target and does not rely on additional assisted mechanisms, 2) a novel dual-plane fitting algorithm designed for sparse ToF data, expanding the applicability of ToF sensors in robotic vision tasks, 3) A new 3-DoF translation calibration theory for ToF-camera pair, which leverages geometric constraints and can potentially be extended to LiDAR-camera calibration. The proposed method offers a practical, low-cost, and accurate calibration solution, facilitating the broader adoption of ToF sensors in robotic and autonomous systems.

II. RELATED WORK

Robotic and autonomous systems are increasingly integrating multiple sensors to enhance accuracy and reliability in complex tasks[10]. Calibration for sensors with different data modalities is typically achieved through time-triggered calibration, which synchronizes data acquisition across multiple sensors[11]. This synchronization can be implemented via software-based[12] or hardware-based[11] triggering mechanisms.

In the case of extrinsic calibration, which involves estimating the 3-degree-of-freedom (3-DoF) rotation and 3-DoF translation between sensors, the key challenge is to identify correspondences between the same real-world features in different sensor coordinate frames. For stereo vision systems, chessboard patterns are widely used as calibration targets[13].

For LiDAR-camera calibration, planar and linear features are commonly utilized[14], [15]. A widely adopted approach

involves enforcing geometric constraints where a point lies on a plane or a line[16]. Zhou et al. explored boundary constraints on calibration boards[15], while Dong et al. leveraged line features from calibration patterns[17]. Peyman et al. introduced a method that uses object edges extracted from 2D images and corresponding 3D point clouds for range sensor and camera calibration[18].

ToF sensors pose additional challenges for calibration due to their inherently low resolution, which makes it difficult to accurately extract point or line correspondences. To address this issue, Li et al. proposed an RGB-D SLAM (simultaneous localization and mapping) -based calibration method for ToF-based fine-depth map regression[6]. Their approach requires sequential reconstruction of three distinct planes in 3D space for extrinsic calibration. Similarly, Carter et al. calibrated a ToF sensor mounted on a robotic arm by leveraging the known kinematics of the arm[5], which can also be seen as an extrinsic calibration method.

For rotation calibration, constraints derived from vanishing points (formed by parallel lines) and surface normal vectors have been widely used in prior research[19]. These projective geometric constraints have been successfully applied to LiDAR-camera calibration[20]. To our knowledge, this is the first work that applies such constraints to ToF-camera calibration, and our proposed translation calibration method is novel for ToF sensors.

III. METHODS

In this section, we first introduce the hardware setup. Next, we present the two-plane fitting algorithm for the ToF sensor, which plays a crucial role in extrinsic calibration. Finally, we describe the vanishing point-based rotation calibration method and the plane intersection-based translation calibration method.

A. Hardware Setup

To facilitate future research, we use the RealSense D435i as the RGB camera and its high-resolution depth map for evaluation. A 3D-printed fixture platform is used (2), enabling design modifications to accommodate other sensors.

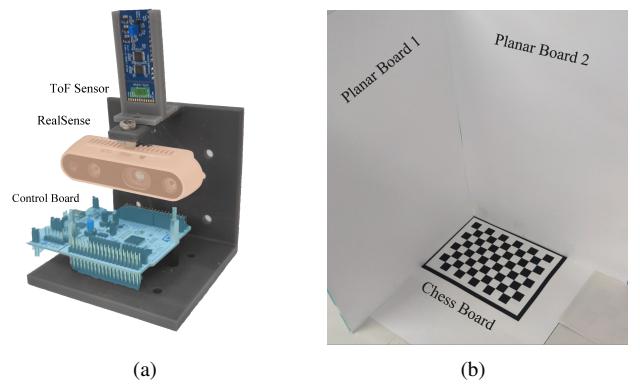


Fig. 2: Hardware setup. (a) The fixture platform for ToF sensor and camera. (b) The calibration board.

¹<https://www.st.com/en/ecosystems/satel-vl53l8.html>

²<https://www.intelrealsense.com/depth-camera-d435i/>

The calibration board consists of three perpendicular planes as shown in Fig. 2, with one featuring a chessboard pattern and the other two being plain white boards.

The chessboard is custom-designed with a black contour, facilitating intersection line detection. In our hardware setup, the RealSense camera is positioned upside down. Since the captured images need to be flipped for observation, we calibrate the flipped camera coordinate system with the ToF sensor's coordinate system and apply a reverse transformation matrix to the final calibration result.

B. ToF Multi-plane Fitting Algorithm

The data from the ToF sensor is sparse, which means it's not point-wise accurate; the zone's value is not the distance for the corresponding point in the real world. The principle of ToF sensor is shown on Fig. 1. As the presupposition in[6], there must at least be one position inside the zone, whose depth is exactly the value of this zone. So ideally, if we up-sample ToF data enough, the precise position of this position can be found. However, it's a waste of time and memory to up-sample the data to a high resolution, and the other up-sampled points where the distance actually is not itself would introduce deviation. So we propose a ToF-

Algorithm 1 ToF-RANSAC Pseudocode

```

Input: ToF data  $\mathcal{D}_{N \times N} \in \mathbb{R}^3$ , Up-sample rate  $T$ 
Output: Two planes  $\mathcal{P}_1(\mathbf{n}_1, d_1), \mathcal{P}_2(\mathbf{n}_2, d_2)$ 
Initialize:  $\mathcal{P}_1, \mathcal{P}_2$ ; RANSAC( $Pro, Threshold, max\_iter$ )
Divide  $\mathcal{D}_{N \times N}$  by  $\nabla \mathcal{D}_{N \times N}$  into  $\mathcal{D}_1, \mathcal{D}_2$ 
for Plane of  $\mathcal{P}_i, i = 1, 2$  do
    while  $iter < max\_iter$  do
        Randomly select 5 pairs offset  $(u_k, v_k), 0 \leq u_k, v_k < T, k = 0, 1, 2, 3, 4$ 
         $\hat{\mathcal{P}}_i(\hat{\mathbf{n}}_i, \hat{d}_i) \leftarrow$  SLSQP optimization to fit  $\mathcal{P}_i$  by minimizing  $\|\mathbf{n}_i \cdot \mathbf{p} + d_i\|$  of  $[x_{ik} + u_k, y_{ik} + v_k, D(x_{ik}, y_{ik})]$ 
        Reproject ToF data to real world scale  $[x_{ik} + u_k, y_{ik} + v_k, D(x_{ik}, y_{ik})]$ 
        for each point  $p[x_{ik}, y_{ik}, D(x_{ik}, y_{ik})]$  in  $\mathcal{D}_i$  do
            Randomly select  $N_i$  pairs offset  $(u_l, v_l), 0 \leq u_l, v_l < T, l = 0, 1, 2, \dots, N_i$ , where  $N_i$  is the number of points in  $\mathcal{D}_i$ 
            Reproject ToF data to real world scale  $[x_{il} + u_l, y_{il} + v_l, D(x_{il}, y_{il})]$ 
            if  $distance(\mathcal{P}_i, [x_{il} + u_l, y_{il} + v_l, D(x_{il}, y_{il})]) < Threshold$  then
                 $inliers_i \leftarrow inliers_i \cap [x_{il}, y_{il}, D(x_{il}, y_{il})]$ 
            end if
             $I_i \leftarrow size(inliers_i)$ 
            if  $I_i$  is enough then
                Output:  $\mathcal{P}_i(\mathbf{n}_i, d_i)$ 
            end if
             $max\_iter \leftarrow \frac{\log(1-Pro)}{\log(1-pow(I_i/N_i, 5))}$ 
             $iter \leftarrow iter + 1$ 
        end for
    end while
end for

```

RANSAC algorithm to achieve a balance between accuracy and efficiency.

ToF-RANSAC firstly randomly selects pairs of offset (u, v) to shift the point's position inside zones. At least, 3 points can fit a plane. We take the setup of FASTER-LIO[21], using 5 points to fit a plane, to increase the robustness. About the fitting algorithm, we utilize SLSQP (Sequential Least SQuare Programming) to optimize the objective function of $\|\mathbf{n}^T \cdot \mathbf{p} + d\|$, which is calculating the distance between point p and plane $\mathcal{P}(\mathbf{n}, d) \in \mathbb{R}^3$. Afterwards, this equation is also applied to calculate error between every points and the fitting plane, to determine the inliers and error, as well as update the maximum iteration. This program would break when the inliers are enough or reach the maximum iteration.

C. Rotation Calibration

To calibrate the rotation, we extract parallel lines from the RGB camera and obtain normal vectors from the ToF sensor as correspondences. The calibration board, shown in Fig. 2, consists of a chessboard pattern and two perpendicular white boards. The normal vectors of the vertical white boards are expected to align with the parallel lines of the chessboard. As a result, their projections onto the image plane should intersect these parallel lines at vanishing points as shown in Fig. 3.

Given the vanishing points $V_i \in \mathbb{R}^2$ corresponding to the width and length of the chessboard image, we convert them into real-world coordinates using:

$$\mathbf{v}_i = \|K^{-1}V_i\| \quad (1)$$

where K is the camera intrinsic matrix. For the ToF data (u, v, d) , we first project it to real-world coordinates (u_w, v_w, d) and apply a ToF-RANSAC plane fitting algorithm to estimate the two planar surfaces: $\mathcal{P}_1(\mathbf{n}_1^T, d_1), \mathcal{P}_2(\mathbf{n}_2^T, d_2)$, where \mathbf{n}_i denote the plane normal vectors. The rotation matrix relates these normal vectors to the vanishing direction vectors, transforming the ToF coordinate system into the camera coordinate system:

$$\mathbf{v}_i = R \cdot \mathbf{n}_i \quad (2)$$

To solve for the rotation matrix R , we apply singular value decomposition (SVD) as described in[22]. Given the



Fig. 3: In the projective geometry, parallel lines will intersect into the vanishing points. The normal vector of the perpendicular plane is also parallel to these lines, so that they should intersect into the same vanishing point. In this figure, normal vector₁ goes to vanishing point₂. On the contrary, normal vector₂ goes to vanishing point₁.

correspondence matrix:

$$S = \sum \mathbf{n}_i \cdot \mathbf{v}_i \quad (3)$$

its SVD decomposition is:

$$\text{SVD}(S) = U\Sigma V^T \quad (4)$$

From which the rotation matrix can be calculated as:

$$R = VU^T \quad (5)$$

To ensure orthogonality of the resulting rotation matrix, we enforce:

$$\begin{cases} \mathbf{r}_1 = \text{norm}(\mathbf{r}_1) \\ \mathbf{r}_2 = \text{norm}(\mathbf{r}_2) \\ \mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2 \end{cases} \quad (6)$$

where \mathbf{r}_i are the columns of the rotation matrix R .

When collecting images, ensure ToF sensor can percept 2 vertical planes, and camera can capture the chessboard. Since each image capture provides two sets of correspondences, at least two different captures with varied orientations are required to fully determine the 3-DoF rotation. To improve robustness, we collect approximately 15~30 image sets from different poses to obtain a more accurate estimation.

D. Translation Calibration

Translation calibration is more difficult than rotation because it's hard to acquire exact position information from sparse ToF data. To address this, we leverage the intersection of planes to obtain the position of point and line indirectly as shown in Fig. 4. Firstly, transfer the ToF fitted planes to camera coordinate system by:

$$\mathbf{n}'_i = R \cdot \mathbf{n}_i, d'_i = d_i - \mathbf{n}'_i^T \cdot \mathbf{t} \quad (7)$$

where \mathbf{t} is the translation vector. The intersection point of three planes can be calculated as:

$$\begin{bmatrix} \mathbf{n}'_1^T & d'_1 \\ \mathbf{n}'_2^T & d'_2 \\ \mathbf{n}'_3^T & d'_3 \end{bmatrix} \mathbf{O} = \mathbf{0} \quad (8)$$

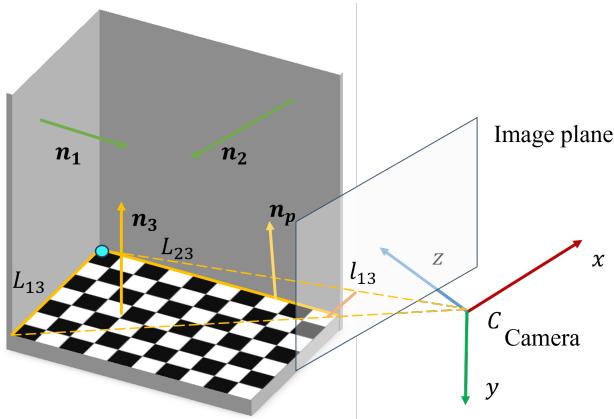


Fig. 4: To distinguish the parameters in different coordinate system, the ToF parameters are in green, while the camera parameters are in orange. The cyan point is the intersection point of 3 planes, and L_{13}, L_{23} are the intersections planes. l_{13} is L_{13} on the image plane.

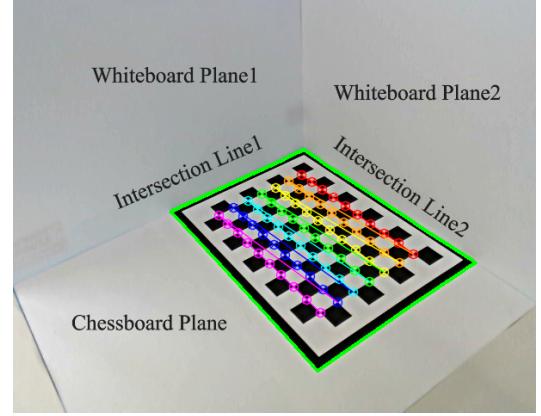


Fig. 5: The RGB camera's view, where colorful corner points are detected for vanishing point and \mathbf{n}_3, d_3 calculation. The green contour is designed for detecting intersection lines between the vertical and horizontal planes.

where \mathbf{O} is the homogeneous intersection point.

The 3 planes constraints have been utilized, new constraints are required to restrict the point \mathbf{O} to solve the unknown \mathbf{t} . Moreover, 3D line is difficult to represent in projective geometry. Because 3D line owns 4-DoF, if record it in a homogeneous way it owns 5 dimensions[23]. So we try to figure out an intuitive way to solve this problem. Notice that \mathbf{O} lays on the intersection lines between the chessboard with the two white boards, so that it also lays on the plane generated by the intersection lines on image plane with the camera optical center. The generated plane can be represented by:

$$\mathcal{P}_n = [K^T \cdot l, 0]^T \quad (9)$$

if the intersection line on the image plane can be represented by $l \in \mathbb{R}^2$. $K^T \cdot l$ can be recorded as the normal vector \mathbf{n}_p . The intersection point $\mathbf{O} \in \mathbb{R}^3$ can be constrained by:

$$[\mathbf{n}_p^T, 0] \mathbf{O} = 0 \quad (10)$$

where only \mathbf{t} is unknown after we have calibrated the rotation R . The close form solution can be found in the Appendixes.

To extract the intersection lines on the image, we design the chessboard with a black contour as shown on Fig.2. The feature extraction result is on Fig.5.

IV. EXPERIMENTS

The ToF sensor used in this study is the STMicroelectronics VL53L8A, a commercially available sensor with a breakout board. The core sensor is compact, measuring $6.4 \times 3.0 \times 1.75$ mm (excluding the breakout board). It operates based on single-photon avalanche diodes (SPADs) and provides an output resolution of either 4×4 or 8×8 pixels. The sensor estimates depth by counting photons received from different distances in each zone and fitting a histogram to determine an expected depth value.

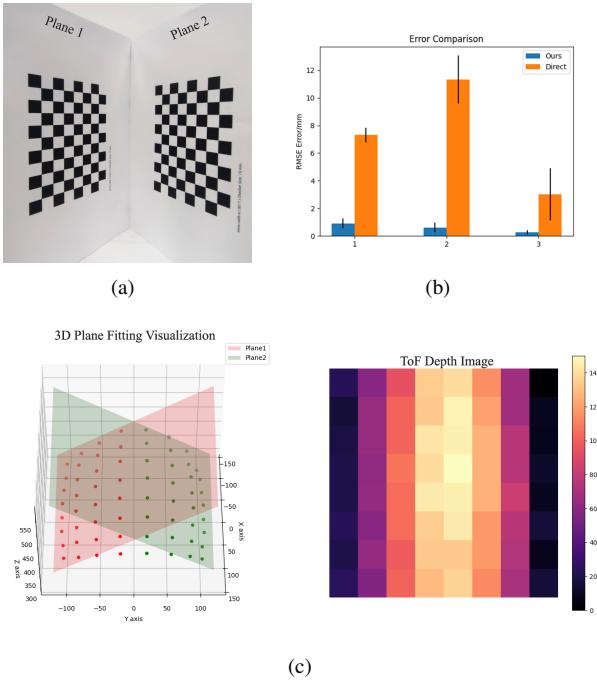


Fig. 6: 2 plan fitting result by ToF-RANSAC. (a) ToF sensor’s view. (b) The mean error comparison of 3 trials. (c) Left is 3D visualization. Right is the depth image of ToF sensor.

A. Plane Fitting

For the plane fitting experiments, our ToF-RANSAC algorithm prioritizes accuracy, which is crucial for calibration tasks. Compared to the conventional approach of directly fitting a plane using all inlier points—referred to as the “Direct Method”—our ToF-RANSAC algorithm achieves a threefold reduction in error, as shown in Fig. 6.

Additionally, due to its randomized selection process, our method effectively mitigates deviation artifacts, as illustrated in Fig. 7. This ensures that the bottom chessboard plane does not affect the vertical plane fitting, provided that the majority of points belong to the vertical planes.

B. Calibration

To evaluate the calibration accuracy, as the ground truth of rotation and translation cannot be obtained, the depth maps before and after calibration are utilized for evaluation. We utilize the fine depth map from the RealSense camera as reference. The first step involves aligning the RealSense depth map with its corresponding RGB image, ensuring that the depth value at each pixel position in the RGB image corresponds precisely to its indexed depth in the depth map. Since the ToF sensor has been calibrated with the RealSense RGB camera, the calibration implicitly establishes a transformation between the ToF sensor and the RealSense depth map, allowing direct comparison.

To account for the resolution disparity between ToF sensor and RealSense, RealSense’s depth map is divided into 8×8 zones, corresponding to ToF sensor’s resolution. Within

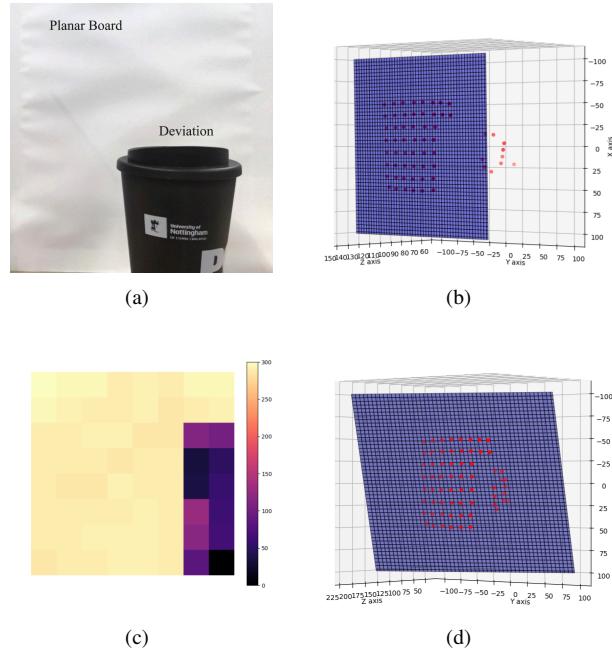


Fig. 7: Robustness Evaluation Experiment. A cup in front of a white board introduces deviation in plane fitting. (a): ToF sensor’s view. (c): Depth data, with purple zones indicating closer deviation. (b): Our algorithm excludes deviation points, fitting the purple plane to red 3D points. (d): The direct method skews the fitted plane due to outliers.

each zone, a depth histogram is computed, and a Gaussian curve is fitted to obtain the expectation value, which serves as the representative depth measurement for that region. The evaluation is then conducted by comparing the ToF sensor’s depth measurements before and after calibration against the processed RealSense depth map.

Furthermore, as the field of view (FoV) of the ToF sensor is square, whereas the RealSense depth map has a resolution of 640×480 , as shown in Fig. 8, an alignment process is applied to match the image center with the ToF data center to ensure a fair evaluation before calibration. This step ensures that any discrepancies in the depth measurements originate from calibration inaccuracies rather than misalignment between the two sensors’ fields of view.

In Fig. 8, the upper images display the raw data obtained from the ToF sensor and RealSense separately. The discrep-

TABLE I: The Calibration Error Result (5 Trials)

Trial No	Mean (mm)		RMS (mm)		Std (mm ²)	
	Before	After	Before	After	Before	After
1	69.81	56.72	90.04	74.53	55.72	47.14
2	83.53	55.94	105.14	66.48	63.16	35.19
3	65.02	56.30	81.11	68.84	47.90	38.50
4	85.85	59.88	110.10	69.47	68.23	34.76
5	72.87	48.55	96.54	59.83	62.53	34.35
Average	75.42	55.48	96.59	67.89	59.51	37.99

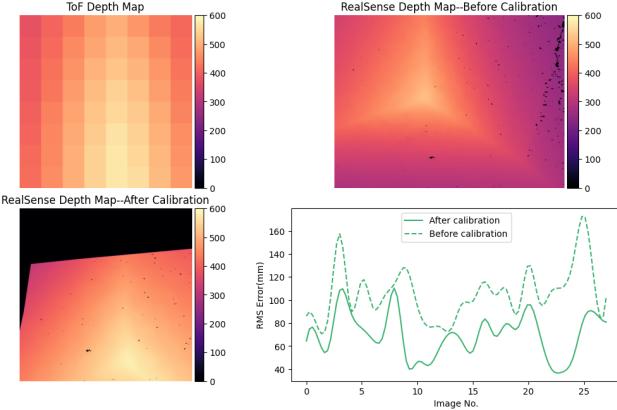


Fig. 8: Depth maps from ToF sensor (*Top-left*) and RealSense before calibration (*Top-right*) and after calibration (*Bottom-left*). We test our algorithm on different perspective images, the RMS error before (dashed) and after calibration is as shown on the bottom-left plot.

ancies in both FoV and depth measurements between the two sensors are evident. After calibration, the RealSense data is aligned to the ToF data, as shown in the bottom-left image, following the evaluation method adopted in this paper. The black regions indicate areas outside the RealSense's FoV. In principle, the alignment can also be performed in the inverse manner, where ToF data is mapped onto the RealSense frame. The mean distances between RealSense's depth map and ToF's data are assessed using the mean and RMS difference over three trials. The average RMS difference of 5 trials average is reduced from 96.59 mm to 67.89 mm as shown in Tab. I, demonstrating the effectiveness of the proposed calibration method.

V. DISCUSSION

This paper proposes an extrinsic calibration method for a ToF sensor and an RGB camera using only three planar boards, eliminating the need for additional assistance such as sensor trajectory reconstruction required in existing approaches. To achieve accurate calibration, a novel dual-plane fitting algorithm is introduced, enhancing both rotational and

translational calibration. Experimental results demonstrate that the proposed algorithm reduces plane fitting error by a factor of three compared to the direct method, decreasing from approximately 7 mm to below 2 mm. Furthermore, the application of the dual-plane fitting algorithm is validated in continuum robot navigation as shown in Fig. 9. This work establishes a foundational framework for ToF-camera extrinsic calibration, providing key considerations and precautions. Future directions include developing a fine-depth estimation model for ToF-camera pairs, leveraging supervision from RealSense, and integrating compact ToF-camera systems into continuum and small-scale robots for advanced applications such as SLAM and 3D reconstruction.

VI. APPENDIX

A. Close Form Solution of Translation Calibration

With the equation (8), rewrite the 3 plane intersection equation as

$$O = \begin{bmatrix} \mathbf{n}_1^T R^T \\ \mathbf{n}_2^T R^T \\ \mathbf{n}_3^T \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{n}_1^T R^T \mathbf{t} - d_1 \\ \mathbf{n}_2^T R^T \mathbf{t} - d_2 \\ -d_3 \end{bmatrix} \quad (11)$$

where O is the inhomogeneous representation. So that equation (10) can be written as:

$$\mathbf{n}_p^T \begin{bmatrix} \mathbf{n}_1^T R^T \\ \mathbf{n}_2^T R^T \\ \mathbf{n}_3^T \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{n}_1^T R^T \mathbf{t} - d_1 \\ \mathbf{n}_2^T R^T \mathbf{t} - d_2 \\ -d_3 \end{bmatrix} = 0 \quad (12)$$

Extract \mathbf{t} :

$$\mathbf{n}_p^T \begin{bmatrix} \mathbf{n}_1^T R^T \\ \mathbf{n}_2^T R^T \\ \mathbf{n}_3^T \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{n}_1^T \\ \mathbf{n}_2^T \\ \mathbf{0}^T \end{bmatrix} \mathbf{t} = \mathbf{n}_p^T \begin{bmatrix} \mathbf{n}_1^T R^T \\ \mathbf{n}_2^T R^T \\ \mathbf{n}_3^T \end{bmatrix}^{-1} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \quad (13)$$

If take:

$$N = \begin{bmatrix} \mathbf{n}_1^T R^T \\ \mathbf{n}_2^T R^T \\ \mathbf{n}_3^T \end{bmatrix}, M = \begin{bmatrix} \mathbf{n}_1^T \\ \mathbf{n}_2^T \\ \mathbf{0}^T \end{bmatrix}, \mathbf{d} = [d_1, d_2, d_3]^T \quad (14)$$

\mathbf{t} can be calculated as:

$$\mathbf{t} = ((\mathbf{n}_p^T \cdot NM)^T (\mathbf{n}_p^T \cdot NM))^{-1} (\mathbf{n}_p^T \cdot NM)^T \mathbf{n}_p^T \cdot N \mathbf{d} \quad (15)$$

B. Calibration Result

The rotation matrix and translation vector for this work's hardware are

$$R = \begin{bmatrix} 0.9812677 & -0.04414566 & -0.18752294 \\ -0.00818207 & 0.96296267 & -0.26951058 \\ 0.19247532 & 0.26599635 & 0.94456296 \end{bmatrix}$$

$$\mathbf{t} = [-45.50863735, -39.29209392, 137.2972859]^T$$



Fig. 9: Application of the Two-Plane Fitting Algorithm. (a) A ToF sensor mounted at the tip of a continuum robot. In an industrial scenario, the robot encounters a fork in the road. (b) The robot selects the more inclined path to maximize its reach.

REFERENCES

- [1] R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," *IEEE Transactions on Robotics*, vol. 33, pp. 1255–1262, Oct. 2017.
- [2] H. Yim, H. Kang, T. D. Nguyen, and H. R. Choi, "Electromagnetic Field & ToF Sensor Fusion for Advanced Perceptual Capability of Robots," *IEEE Robotics and Automation Letters*, vol. 9, pp. 4846–4853, May 2024.
- [3] A. Marcu, D. Costea, V. Licăreț, M. Pîrvu, E. Slușanschi, and M. Leordeanu, "SafeUAV: Learning to Estimate Depth and Safe Landing Areas for UAVs from Synthetic Data," in *Computer Vision – ECCV 2018 Workshops* (L. Leal-Taixé and S. Roth, eds.), vol. 11130, pp. 43–58, Cham: Springer International Publishing, 2019.
- [4] C. Sifferman, Y. Wang, M. Gupta, and M. Gleicher, "Unlocking the Performance of Proximity Sensors by Utilizing Transient Histograms," Aug. 2023.
- [5] C. Sifferman, W. Sun, M. Gupta, and M. Gleicher, "Using a Distance Sensor to Detect Deviations in a Planar Surface," Aug. 2024.
- [6] Y. Li, X. Liu, W. Dong, H. Zhou, H. Bao, G. Zhang, Y. Zhang, and Z. Cui, "DELTAR: Depth Estimation from a Light-Weight ToF Sensor and RGB Image," in *Computer Vision – ECCV 2022* (S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner, eds.), vol. 13661, pp. 619–636, Cham: Springer Nature Switzerland, 2022.
- [7] F. Mu, C. Sifferman, S. Jungerman, Y. Li, M. Han, M. Gleicher, M. Gupta, and Y. Li, "Towards 3D Vision with Low-Cost Single-Photon Cameras," Mar. 2024.
- [8] X. Liu, Y. Li, Y. Teng, H. Bao, G. Zhang, Y. Zhang, and Z. Cui, "Multi-Modal Neural Radiance Field for Monocular Dense SLAM with a Light-Weight ToF Sensor," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, (Paris, France), pp. 1–11, IEEE, Oct. 2023.
- [9] C. Sifferman, D. Mehrotra, M. Gupta, and M. Gleicher, "Geometric Calibration of Single-Pixel Distance Sensors," *IEEE Robotics and Automation Letters*, vol. 7, pp. 6598–6605, July 2022.
- [10] M. B. Alatise and G. P. Hancse, "A Review on Challenges of Autonomous Mobile Robot and Sensor Fusion Methods," *IEEE Access*, vol. 8, pp. 39830–39846, 2020.
- [11] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *The International Journal of Robotics Research*, vol. 32, pp. 1231–1237, Sept. 2013.
- [12] H. Hu, J. Wu, and Z. Xiong, "A Soft Time Synchronization Framework for Multi-Sensors in Autonomous Localization and Navigation," in *2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, (Auckland), pp. 694–699, IEEE, July 2018.
- [13] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330–1334, Nov. 2000.
- [14] A. Geiger, F. Moosmann, O. Car, and B. Schuster, "Automatic camera and range sensor calibration using a single shot," in *2012 IEEE International Conference on Robotics and Automation*, (St Paul, MN, USA), pp. 3936–3943, IEEE, May 2012.
- [15] L. Zhou, Z. Li, and M. Kaess, "Automatic Extrinsic Calibration of a Camera and a 3D LiDAR Using Line and Plane Correspondences," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Madrid), pp. 5562–5569, IEEE, Oct. 2018.
- [16] Z. Pusztai and L. Hajder, "Accurate Calibration of LiDAR-Camera Systems Using Ordinary Boxes," in *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, (Venice, Italy), pp. 394–402, IEEE, Oct. 2017.
- [17] W. Dong and V. Isler, "A Novel Method for the Extrinsic Calibration of a 2D Laser Rangefinder and a Camera," *IEEE Sensors Journal*, vol. 18, pp. 4200–4211, May 2018.
- [18] P. Moghadam, M. Bosse, and R. Zlot, "Line-based extrinsic calibration of range and image sensors," in *2013 IEEE International Conference on Robotics and Automation*, (Karlsruhe, Germany), pp. 3685–3691, IEEE, May 2013.
- [19] B. Caprile and V. Torre, "Using vanishing points for camera calibration," *International Journal of Computer Vision*, vol. 4, pp. 127–139, Mar. 1990.
- [20] Y. Zhao, L. Zhao, F. Yang, W. Li, and Y. Sun, "Extrinsic Calibration of High Resolution LiDAR and Camera Based on Vanishing Points," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–12, 2024.
- [21] C. Bai, T. Xiao, Y. Chen, H. Wang, F. Zhang, and X. Gao, "Faster-LIO: Lightweight Tightly Coupled Lidar-Inertial Odometry Using Parallel Sparse Incremental Voxels," *IEEE Robotics and Automation Letters*, vol. 7, pp. 4861–4868, Apr. 2022.
- [22] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-Squares Fitting of Two 3-D Point Sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, pp. 698–700, Sept. 1987.
- [23] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2 ed., Mar. 2004.