

SPLiCE: Single-Point LiDAR and Camera Calibration & Estimation Leveraging Manhattan World

Minji Kim, Jeahn Han, Jungil Ham, and Pyojin Kim

Abstract—We present a novel calibration method between single-point LiDAR and camera sensors utilizing an easy-to-build customized calibration board satisfying the Manhattan world (MW). Previous methods for LiDAR-camera (LC) calibration focus on line and plane correspondences. However, they require dense 3D point clouds from heavy and expensive LiDAR to simplify alignments; otherwise, these approaches fail for extremely sparse LiDAR. Compact, lightweight, and sparse LiDAR and camera sensors are inevitable for micro drones like Crazyflie with a maximum payload of 15 g, but there are no explicit calibration methods for them. To address these issues, we propose a new extrinsic calibration method with a new calibration board, which rotates like a door to capture geometric features and align them with images. Once we find an initial estimate, we refine the relative rotation by minimizing the angle difference between the grid orientation of the checkerboard and the MW axes. We demonstrate the effectiveness of the proposed method through various LC configurations, achieving its capability and high accuracy compared to other state-of-the-art approaches. We release our calibration toolkit, source codes, and how to make the calibration boards for the robotics community: <https://SPLiCE-Calib.github.io/>.

I. INTRODUCTION

Extrinsic calibration between LiDAR and camera sensors is one of the fundamental building blocks for various sensor fusion algorithms and perception tasks, widely used in robotics, autonomous driving, and computer vision applications. Although extensive research has focused on LiDAR-camera (LC) extrinsic calibration [1], [2], [3] so far, most existing methods are targeting bulky multi-beam LiDAR sensors (Velodyne VLP-16, Ouster OS1-64, and Livox Avia) that produce dense point clouds. These high-resolution and dense point clouds simplify extrinsic calibration tasks thanks to their rich spatial information, allowing for robust feature matching and geometric alignments. However, such LiDAR sensors are often heavy, expensive, and energy-intensive, making them unsuitable for compact, lightweight robotic platforms with limited resources.

The need for extremely sparse LiDAR sensors has grown with their expanding use, such as nano drone navigation [4], [5], wearable systems [6], motion estimation [7], [8], [9], and depth completion tasks [10], [11]. These single-point LiDAR sensors, although significantly constrained in data density, can offer unique advantages in terms of size, weight, and

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.RS-2024-00358374). (Corresponding author: Pyojin Kim)

Department of Mechanical and Robotics Engineering, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, South Korea. {minji0110, jeahnhaan, jungilham}@gm.gist.ac.kr, {pjinkim}@gist.ac.kr

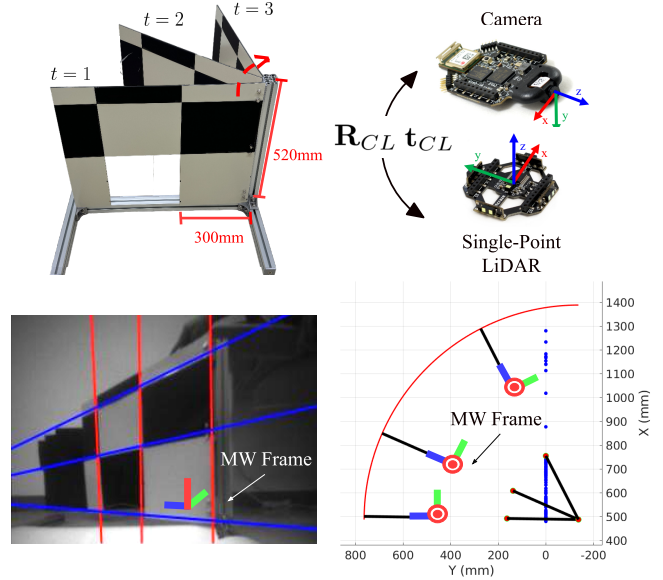


Fig. 1. Proposed custom-built calibration tool with checkerboard that rotates like a door (top-left) for single-point LiDAR and camera extrinsic calibration (top-right). We can achieve accurate spatial alignments between them with lines and vanishing points in the images (bottom-left) satisfying Manhattan world (MW), and the accumulated points from single-point LiDAR (bottom-right).

power consumption, making them ideal for integration into various robotic and computer vision applications. This lack of spatial density, however, poses significant challenges for LiDAR-camera (LC) extrinsic calibration because traditional methods heavily rely on rich and dense point clouds per frame from LiDAR sensors.

To address these issues, we propose a novel extrinsic calibration method between an extremely sparse single-point LiDAR and a camera, which exploits point and line features within a custom-built checkerboard satisfying MW as shown in 1. Our method leverages the structural geometric properties of an easy-to-build customized calibration tool that rotates like a door from a camera and a single-point LiDAR, and spatiotemporally aligns the two sensors. Through intra-drone and inter-drone single-point LiDAR-camera calibration experiments across various environments, we demonstrate the robustness and accuracy of the proposed method, showing its potential and capability for lightweight platforms like a Crazyflie [12] where classical dense LiDAR sensors are unsuitable. Our main contributions are as follows:

- We design an easy-to-build calibration target board with a checkerboard, which rotates like a door to overcome the extreme data sparsity of a single-point LiDAR.
- We propose a novel extrinsic calibration method that

leverages the structural regularities of the calibration tool with points and lines obtained by rotating the target board.

- We validate our approach through extensive LiDAR and camera configurations on the Crazyflie nano drone, demonstrating its capability and accuracy compared to the existing methods.

II. RELATED WORK

Traditional LC Calibration. Estimating the rigid-body transformation for LiDAR-camera (LC) extrinsic calibration is well-studied, with various solutions that can be categorized into target-based [1], [13], [2], [14], target-less [15], [16], [17], [18], and learning-based [19], [20], [3], [21] methods. Target-based methods require pre-defined calibration objects that provide explicit geometric features while target-less and learning-based methods leverage natural features from the surrounding environments rather than depending on specific targets. They require rich and dense point clouds per frame from heavy, bulky, and expensive LiDAR sensors, allowing for robust feature matching and geometric alignments between them.

Target-based Single-Point LC Calibration. Some extrinsic calibration methods for single-point LiDAR employ geometric constraints to overcome the extreme sparsity of the LiDAR data. Compared to heavy and dense LiDARs, calibrating the relative pose between a single-point LiDAR and a camera requires more constraints due to the limited data from a single-point measurement. [22] proposes two methods using a calibration board printed with circles. They require either the visible 2D coordinates of the laser dot on the target board or the spatial pre-defined relationship between the camera and the target plane. [23] utilizes spherical objects as targets, applying point-to-line geometric constraints to achieve precise alignment, improving calibration accuracy by minimizing the projection error of the detected points along specific projected lines. [24] handles invisible laser spots using constraint equations, assuming known relative poses between infrared and main cameras. [25] establishes a correspondence between range measurement distances and the pixel coordinates of laser spots on the imaging plane. [26] proposes a framework when the visibility of the laser spot is unclear or inaccurate. By utilizing the local image intensity of the laser spot boundary, enabling high-accuracy pixel coordinates even under challenging conditions. Existing methods rely on visible laser points, making them sensitive to lighting changes.

Motion-based Single-Point LC Calibration. There are some calibration approaches to overcome the extreme data sparsity by accumulating a single-point over time through estimated the six degrees of freedom (DoF) trajectories with the help of odometry and SLAM. [27] proposes a method of estimating LiDAR motion through LiDAR SLAM and camera motion via visual or fusion SLAM. This method works without a target but depends on LiDAR odometry and SLAM accuracy. If 1D data is extended to 2D or 3D, calibration methods using dense data for 2D or 3D LiDAR

can be applied to 1D. For example, some methods find correspondences by using the edges and right-angle structures of the target [28], [29]. Additionally, other methods use target boards with holes or transparent areas to match 2D and 3D coordinates [30], [31], [32], while target-less approaches exist for reconstructing 3D structures and aligning them with a 3D point cloud [14], [33]. Motion-based single-point LC calibration methods have the limitation of requiring precise and accurate odometry and SLAM results. The proposed method addresses these limitations by introducing a new calibration target board and leveraging structural regularities of the MW model, where all lines are mutually orthogonal, to overcome these challenges.

III. PRELIMINARY

A. Notations

The 2D pixel coordinates (u, v) are obtained by projecting a 3D point (X_C, Y_C, Z_C) in the camera frame using the intrinsic matrix. The single-point LiDAR provides only distance values, so in the LiDAR frame (Fig. 1), the data lies on the xy -plane. We can represent the relative 6-DoF pose between the camera and the single-point LiDAR by following Eq.(1) and Eq.(2).

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \mathbf{T}_{CL} \begin{bmatrix} X_L \\ Y_L \\ 0 \\ 1 \end{bmatrix}, (\because P_L = (X_L, Y_L, 0)^\top) \quad (1)$$

$$\mathbf{T}_{CL} = \begin{bmatrix} \mathbf{R}_{CL} & \mathbf{t}_{CL} \\ \mathbf{0}^\top & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix} \in \text{SE}(3) \quad (2)$$

The transformation matrix \mathbf{T}_{CL} consists of rotation $\mathbf{R}_{CL} \in \text{SO}(3)$ and translation $\mathbf{t}_{CL} \in \mathbb{R}^3$.

$$\tilde{P}_L = \begin{bmatrix} u_L \\ v_L \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X_L \\ Y_L \\ 1 \end{bmatrix} = \mathbf{K}\mathbf{E} \begin{bmatrix} X_L \\ Y_L \\ 1 \end{bmatrix} \quad (3)$$

where \tilde{P}_L is the projection of $P_L = (X_L, Y_L, Z_L)^\top$ from the LiDAR frame onto the image plane as given by Eq.(3). Because $Z_L = 0$, the remaining components $[\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{t}]$ form \mathbf{E} , the extrinsic parameters between the LiDAR and camera frames.

B. Gaussian Sphere

A Gaussian sphere is a unit sphere centered at the camera's projection center. As shown in Fig. 2, lines detected in the image, when projected onto this sphere, appear as great circles. Parallel lines in the images form great circles that intersect at two antipodal points, and the intersection vector eventually becomes a vanishing direction (VD). When every line and plane is perpendicular to a coordinate axis, we call this configuration the Manhattan world (MW) [34] with the VDs representing the orientation of the Manhattan world frame (MF). The surface normals of orthogonal planes satisfying the MW, which are mutually perpendicular, align with the three VDs from the parallel lines. For full details of the MF on the Gaussian sphere, refer to [35]. The supplementary material provides a summary of all acronyms.

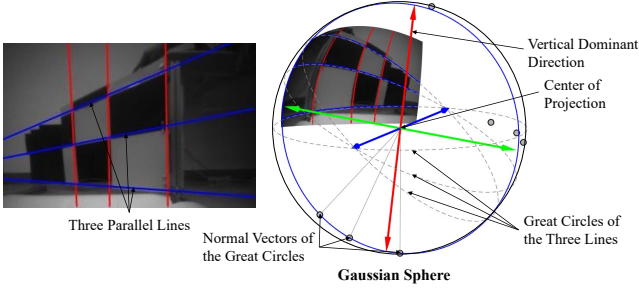


Fig. 2. Geometric relationship between the parallel lines and MW on the Gaussian sphere. The parallel red and blue lines in the image correspond to the red and blue Manhattan frame (MF) directions, respectively.

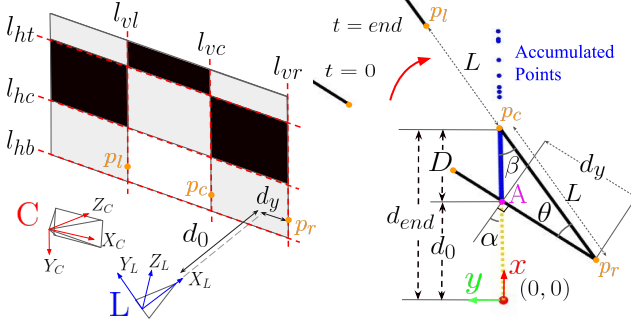


Fig. 3. C and L represent the camera (red) and LiDAR (blue) frames. The camera provides three vertical and horizontal lines, while three orange points are estimated from accumulated single-point LiDAR data (left). Black lines (right) denote the top view of the calibration board with a central hole. The definitions and geometric meanings of each symbol are detailed in Table I.

IV. METHOD

We propose a new calibration method called SPLiCE (Single-Point LiDAR and Camera Calibration) for estimating the extrinsic parameters E in III-A. Our key idea is to establish 2D-3D correspondences between the LiDAR and camera frames by inducing rotational motion on a new calibration target board. Furthermore, we refine the relative rotation between two sensors by aligning the orientation of the calibration target board with respect to the MW axes. Fig. 4 provides an overview of the proposed SPLiCE.

A. Calibration Target Board

We propose a new calibration target board that can effectively extract features in both image and LiDAR distance data. The calibration board is completely flat and includes multiple black-and-white grids as shown in Fig. 1. Each grid row extends 300 mm horizontally, with a square hole at the lower center of the grid. Fig. 3 shows that it can rotate 90 degrees around the rightmost vertical line (l_{vr}), which serves as its axis of rotation with a red arc tracing the rotation path of the board's endpoint as shown in Fig. 1. We assume that the rotation axis of the calibration board and the z -axis of the LiDAR frame should be parallel.

The single-point LiDAR provides only a distance value (d_t) in one direction. We utilize the square hole in the target board to extend the distance value into three points, requiring knowledge of both the grid and hole width. Details are in IV-B. In the image from the camera, as shown at the bottom

TABLE I
DEFINITION OF KEY SYMBOLS

Symbol	Definition
d_0	Distance when LiDAR first hits the calibration board
d_{end}	Distance when LiDAR last hits the calibration board
d_y	Distance between points A and p_r
θ	Rotation angle of the calib. board from $t = 0$ to end
α	Incidence angle of LiDAR to the calib. board at $t = 0$
L	Width of the square grid and hole

left of Fig. 1, it perceives the square hole as a black grid due to the black wall visible through it and does not recognize it as an open hole.

B. Feature Extraction

Three Points Estimation. The distance at $t = 0$ is denoted as d_0 , and d_{end} is the distance measured at the final detection time of the target board by the LiDAR through the square hole as shown in Fig. 3. We keep the target board stationary for over 20 seconds at $t = 0$ and use the average of the measurements during this period to minimize the impact of the noise in low-cost and sparse LiDAR measurements. d_i is the distance at $t = i$, and d_{end} is where Eq. (4) fails, marking a significant distance change.

$$\left| \frac{|d_i - d_{i-1}| - |d_{i-1} - d_{i-2}|}{|d_{i+1} - d_i| - |d_i - d_{i-1}|} - 1 \right| > \epsilon \quad (4)$$

where ϵ serves as the threshold and is empirically set to 0.05 m. After rotating at least three times, we utilize the average value at d_{end} .

As shown in Fig. 2, we obtain the green direction of the MW frame by the detected parallel lines on the board in the image. By comparing the MW frame at $t = 0$ and $t = end$, we determine the rotation angle θ of the calibration board. Given $D = d_{end} - d_0$, L , and θ (see Fig. 3), we can also calculate the incidence angle α , which is a compensation factor to mitigate the assumption that the x -axis of the LiDAR frame and the calibration board should be perpendicular. We define $\beta = 90^\circ - \alpha - \theta$ and its valid range is $\beta \in [0^\circ, 90^\circ - \frac{\theta}{2}]$ due to its geometric properties. We can obtain α by solving the following equation from the laws of sines and cosines as follows:

$$\theta = \arcsin \frac{D \sin \beta}{\sqrt{L^2 + D^2 - 2LD \cos \beta}} \quad (5)$$

where L , D , and θ are all known values (for full details and derivations, refer to the supplementary materials).

We compute d_y as $d_y = D \sin \beta / \sin \theta$ to obtain the position of the three-point features (p_r, p_c, p_l) for each i -th frame as follows:

$$\begin{aligned} p_r^i &= (d_0 - d_y \sin \alpha, -d_y \cos \alpha) \\ p_c^i &= (d_0 - d_y \sin \alpha + L \sin \theta_i, -d_y \cos \alpha + L \cos \theta_i) \\ p_l^i &= (d_0 - d_y \sin \alpha + 2L \sin \theta_i, -d_y \cos \alpha + 2L \cos \theta_i) \end{aligned} \quad (6)$$

2D Line Extraction. We follow the next two steps to detect lines on the calibration target board reliably. First, we extract all candidate lines with both LSD [36] and Hough transform. We filter the detected lines to retain lines identified by both

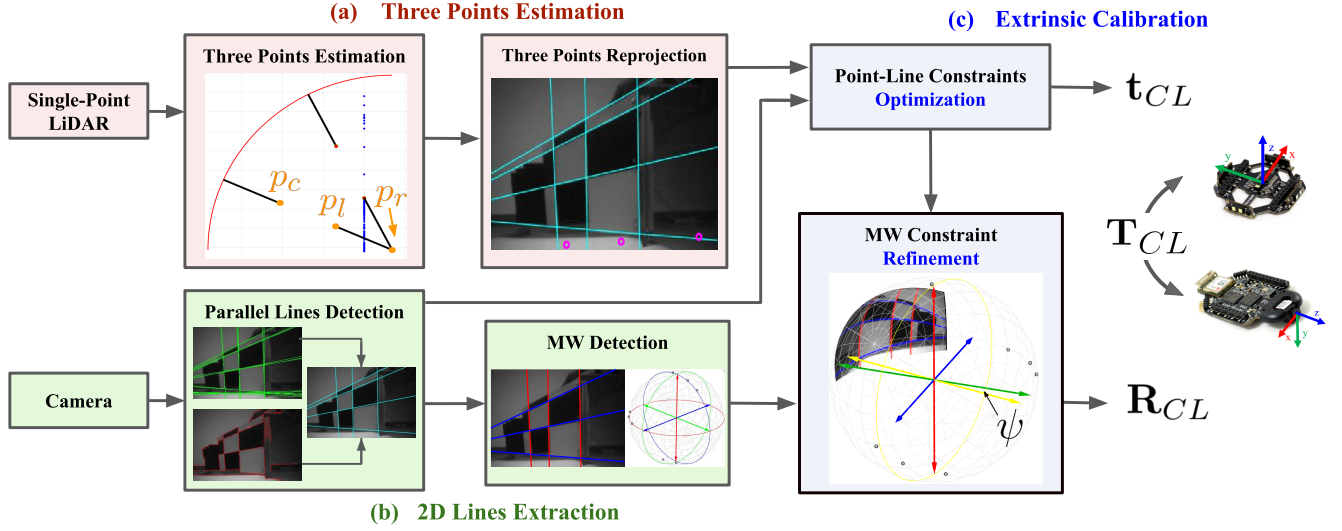


Fig. 4. Overview of the proposed single-point LC extrinsic calibration method. (a) We detect three points (red) based on the accumulated points (blue) from single-point LiDAR by utilizing the geometric characteristics of the proposed new calibration target board. (b) We extract three parallel vertical and horizontal lines in the image and detect the MW aligned to them. (c) Given three points, parallel lines, and the MW, we find the relative rotation and translation between two sensors by minimizing the point-to-line distances and angle differences.

LSD and Hough Transform. A line is considered the same in both methods if its endpoints differ by less than 1.5 pixels and its slope angle (ϕ) deviates by less than 1.5 degrees. We obtain six reliable lines $l^i = (l_{vr}^i, l_{vc}^i, l_{vl}^i, l_{ht}^i, l_{hc}^i, l_{hb}^i)$ from each i -th frame in Eq. (7). If the detected lines from the metal frame or wall remain after filtering, we remove them by discarding those with inter-frame differences below a threshold δ close to zero.

$$l^i = \left\{ \text{line}^i \mid (\| \mathbf{p}_{\text{end}}^{\text{Hough}} - \mathbf{p}_{\text{end}}^{\text{LSD}} \| < 1.5 \text{ pixel}) \cap (|\phi_{\text{Hough}} - \phi_{\text{LSD}}| < 1.5^\circ) \cap (\| \mathbf{p}_{\text{end}}^i - \mathbf{p}_{\text{end}}^{i-1} \| > \delta) \right\} \quad (7)$$

C. Point-Line Constraints Optimization

We obtain the three points $P_L = (p_l, p_c, p_r)$ and project them onto the image plane in Eq.(3), representing as magenta dots in Fig. 5. We categorize the six lines into two groups, three horizontal $l_h = (l_{ht}, l_{hc}, l_{hr})$ and three vertical lines $l_v = (l_{vl}, l_{vr}, l_{vc})$ based on their slopes to establish constraints between each group and the re-projected points $\tilde{p}_L = (\tilde{p}_l, \tilde{p}_c, \tilde{p}_r)$.

First, \tilde{p}_L should lie on the l_v according to the geometric properties of the proposed calibration target board. In the image plane, we measure the Euclidean distance (d_l, d_c, d_r) between each line $l_v = (l_{vl}, l_{vr}, l_{vc})$ and corresponding points $\tilde{p}_L = (\tilde{p}_l, \tilde{p}_c, \tilde{p}_r)$. The cost function consists of the squared sum of point-to-line distances. Second, we can generate a line that extends from \tilde{p}_L , and obtain three intersection points (p_1, p_2, p_3) with the line l_v , which should correspond to the vanishing point formed by the horizontal lines. Originally, the vanishing points of the three lines should coincide, but due to low resolution and pixel noise, the intersections do not align. We obtain three separate intersection points. In the each frame i , we define the coordinates of the intersection points as $\{p_j^i \mid j \in \{1, 2, 3\}\}$. Therefore, we determine the

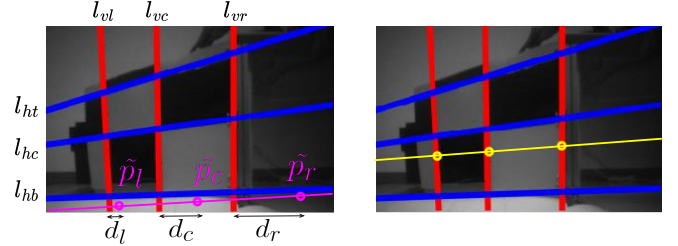


Fig. 5. Three points from LiDAR are re-projected to the image plane (magenta) at the wrong locations with an initial relative pose (left). After optimizing the point-line constraints (right), they are on the three vertical lines well (yellow), and the line connecting the three points is parallel to the horizontal lines.

vanishing point using Eq. (8).

$$\mathbf{p}_c^i = \left(\frac{1}{3} \sum_{j=1}^3 u_j^i, \frac{1}{3} \sum_{j=1}^3 v_j^i \right) \quad (8)$$

We add a constraint that requires (p_1, p_2, p_3) to be the same as p_c . Thus, the distance between the two vanishing points is added to the cost function. We estimate the extrinsic parameters E by solving the following nonlinear problem Eq.(9) using the Levenberg-Marquardt (LM) method:

$$\mathbf{R}^*, \mathbf{t}^* = \arg \min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^N \left(\sum_{k=1}^3 (d_k^i)^2 + \sum_{j=1}^3 (p_c^i - p_j^i)^2 \right) \quad (9)$$

where N is the number of paired images and LiDAR scans.

D. Manhattan World Refinement

Manhattan World Detection. Since the calibration target board is perfectly flat, all lines on the target board form a Manhattan world, so each frame has a different Manhattan frame (MF). We employ 3-line RANSAC [37] to find the current MF in Fig. 6. The sampling-based method produces three VPs, but the exact orientation remains unknown, requiring us to choose one specific case from six possibilities.

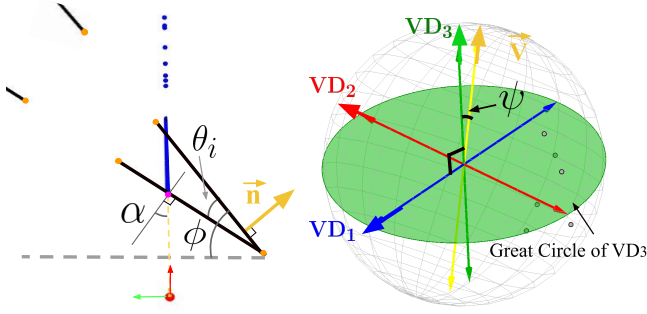


Fig. 6. Top view of the calibration target board in the LiDAR frame (left). \vec{n} represents the normal of the target board expressed in the LiDAR frame, and θ_i and α are from Eq.(6). \vec{V} is \vec{n} expressed in the camera frame (right) given \mathbf{R}^* from Eq.(9). \vec{V} should be parallel to one of the MW axes.

Since we already classify the vertical and horizontal lines, we can leverage this to estimate each approximate VD and denote this frame as the rotation matrix closest to the true orientation, R_t . We define the Manhattan frame (MF) as the set of candidate rotational matrices R_k (where $k = 1, \dots, 6$) that have a rotation error δ with respect to R_t close to either 0 or 180 degrees. We can determine the MF as a single frame based on the rotation error δ by Eq.(10).

$$\delta = \arg \min_{\delta_k} (\min(|\delta_k|, |180^\circ - \delta_k|)) \quad (10)$$

where we validate MW-based tracking accuracy with OptiTrack (for full details, refer to the supplementary materials). **Manhattan World Refinement.** We further refine the relative 3-DoF rotation between two sensors for better accuracy. In Fig. 6, the normal vector \vec{n} of the calibration target board expressed in the LiDAR frame can be transformed to the \vec{V} expressed in the camera frame given the relative rotation:

$$\mathbf{n} = \begin{bmatrix} \cos \phi \\ \sin \phi \\ 0 \end{bmatrix}, \mathbf{R}^* = [\mathbf{r1} \quad \mathbf{r2} \quad \mathbf{r3}] \in \text{SO}(3) \quad (11)$$

$$\mathbf{V} = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = \mathbf{R}^* \begin{bmatrix} \cos \phi \\ \sin \phi \\ 0 \end{bmatrix} \quad (12)$$

ϕ can be expressed as $\theta_i + \alpha$, where Eq. (6) provides θ_i and α . Since \vec{V} should be parallel to the direction of \vec{VD}_3 , which is one of the MW axes in the camera frame, and thus the angle ψ between them should be zero. We find the optimal relative rotation by minimizing the following cost function Eq.(13) using the Levenberg-Marquardt (LM) algorithm:

$$\mathbf{R}^{**} = \arg \min_R \sum_{i=1}^N \left(\frac{\mathbf{V}^i}{\|\mathbf{V}^i\|} \cdot \vec{VD}_3 - 1 \right)^2 \quad (13)$$

where N is the number of paired image and LiDAR scans, and the final extrinsic parameters \mathbf{E} consist of \mathbf{t}_{CL} from Eq.(9) and \mathbf{R}_{CL} from Eq.(13). Note that we use \mathbf{R}^* from the previous step as an initial value for the MW refinement.

V. EVALUATION

A. Calibration Setup

We employ the Crazyflie 2.1 [12] as a micro-robotic platform, which is $9.2 \text{ mm} \times 9.2 \text{ mm}$, weighs 27 g, and

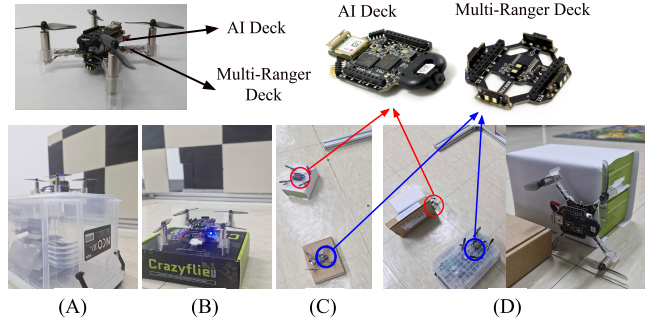


Fig. 7. Four different configurations of single-point LiDAR (multi-ranger deck) and camera (AI deck) sensors. In A and B, both sensors are mounted on the same Crazyflie (intra-drone), but at different heights, while in C and D, two Crazyflies are equipped with either the AI deck or the multi-ranger deck (inter-drone).

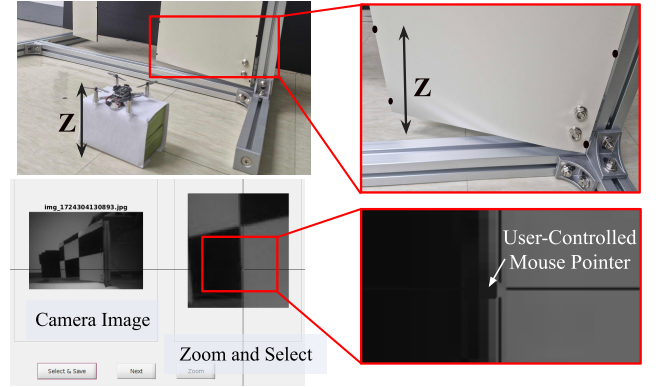


Fig. 8. We need to measure the height Z between the single-point LiDAR and the ground. We mark a black circle on the corresponding point, and identify these three points by zooming in on each image frame.

has a maximum payload of 15 g, only enabling compact and lightweight sensors like the multi-ranger deck equipped with VL53L1x ToF sensors in four directions for distance measurements. We obtain the forward-facing distance values from the multi-ranger deck at 10 Hz to generate a point cloud and continuously accumulate each data point. We communicate with Crazyflie using a laptop connected through the Crazyradio 2.0. We obtain the image sequences from the AI deck equipped with an HM01B0 monochrome camera, with a resolution of 324×244 pixels. The AI deck communicates with the laptop via WiFi, and each image has its own timestamp to be synchronized with Crazyflie's internal clock. Fig.7 shows the multi-ranger deck, AI deck, and various sensor configurations. We calibrate the camera's intrinsic parameters in advance and manually measure relative rotation and translation between the multi-ranger deck and AI deck for an initial guess during the nonlinear optimization in the proposed SPLICE.

We evaluate four different configurations: intra-drone (A and B) and inter-drone (C and D) cases in Fig. 7. We place the multi-ranger deck 17 and 9 cm above the ground in A and B, respectively. To evaluate robustness, significant translational and rotational misalignments between the LiDAR and the camera are introduced in inter-drone setups (C and D). To this end, we mount the AI deck and multi-ranger deck

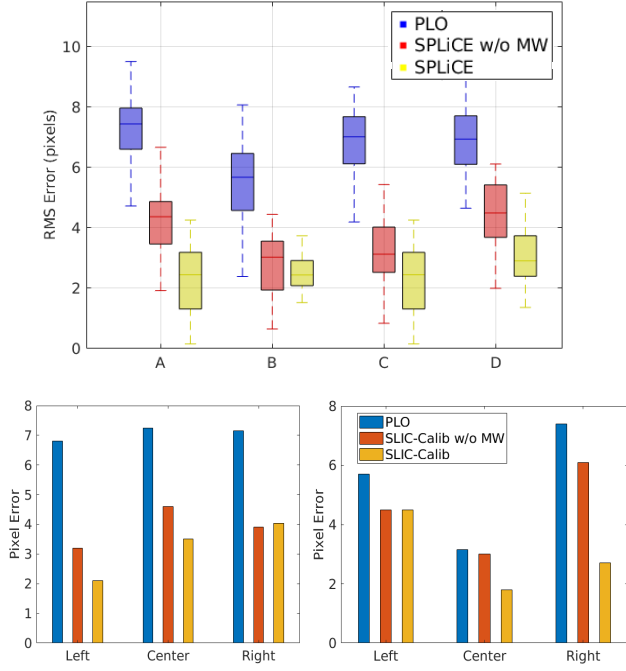


Fig. 9. (Top) Reprojection error distribution across the four configurations: A, B, C, and D. (Bottom) Pixel error statistics of the three points (left, center, right). Our SPLiCE with MW achieves the most accurate results for all configurations and points.

on separate Crazyflies, positioned apart to simulate diverse scenarios. In particular, we rotate the AI deck 90° around the z -axis of the camera frame. In both C and D cases, the multi-ranger deck is positioned 17 cm above the ground.

B. Reprojection Error

We first evaluate the accuracy of extrinsic calibration methods with respect to the reprojection error metric in pixels. We re-project the three points from the single-point LiDAR onto the image plane, and compute the distances between the reprojected and true points on the calibration target board. We measure the height of Crazyflie to the ground and mark the points on the target board, selecting them as ground-truth points for all frames as shown in Fig. 8.

TABLE II
REPROJECTION ERROR RESULTS [PIXEL]

	Intra-Drone		Inter-Drone	
	A	B	C	D
PLO [38]	7.13	5.62	6.91	6.69
SPLiCE w/o MW	4.36	2.91	3.34	4.53
SPLiCE	3.24	2.55	2.59	2.95

The baseline [38] estimates extrinsic parameters between 2D dense LiDAR and the camera by minimizing the point and line distance error. [38] utilizes only three points and vertical lines, so we implement and customize it to our single-point LiDAR data, referred to as PLO (Point-Line Optimization). We also include the calibration results with and without performing the MW refinement step to observe

TABLE III

RELATIVE TRANSLATION AND ROTATION RESULTS IN INTER-DRONE

Inter-Drone	tx [m]	ty [m]	tz [m]	roll [deg]	pitch [deg]	yaw [deg]
C						
SPLiCE w/o MW	0.0098	0.0110	0.0135	1.8344	1.9867	2.0475
SPLiCE	0.0098	0.0110	0.0135	1.1639	1.3869	1.2061
MiniVO	0.0253	0.0314	0.0288	-	6.3924	8.8846
D						
SPLiCE w/o MW	0.0149	0.0173	0.0187	2.0433	1.9001	2.7163
SPLiCE	0.0149	0.0173	0.0187	1.2314	1.0983	1.1863
MiniVO	0.0307	0.0696	0.0471	-	7.1024	9.4527

the effectiveness of the step. Table II shows our comparison results by measuring the RMS pixel distance relative to the ground-truth. The proposed SPLiCE demonstrates an accuracy of 2.83 pixels on average.

Fig. 10 shows the reprojection results of three points from single-point LiDAR onto the image plane with the proposed SPLiCE. It includes four image frames for each configuration, capturing the movements of the calibration target board over time from left to right. In each frame, the green points are consistently closer to the white ground truth points than the yellow points. We can also observe the effectiveness of the MW refinement step in the proposed method, improving overall accuracy.

We also investigate the reprojection error in terms of the number of LiDAR and camera pairs. The number of pairs is randomly chosen from 24 pairs. As shown in Fig. 11, the error value is reduced as the number of pairs increases. With more than 20 pairs, the results improve by over five times compared to using fewer than 10 pairs.

C. Ground Truth Comparison

In experiments with two Crazyflies, each equipped with a multi-ranger deck and an AI deck, as in inter-drone C, D shown in Fig. 7, we attach artificial markers to enable comparison with the ground truth from the OptiTrack motion capture system. We compare the extrinsic calibration results with the ground truth errors, as shown in Table III. The translation error remains the same, owing to the additional refinement applied only to the relative rotation. We achieve an error level of approximately 0.01 m, with rotation errors within about 1 degree. The low-resolution images from the AI deck reduce the precision of lines when using pixel vanishing points. However, the sampling-based RANSAC algorithm improves accuracy by identifying an MW closer to the actual one, producing a more accurate rotation matrix for the extrinsic parameters.

D. Comparison with Other Methods

We evaluate the proposed SPLiCE against two other methods: the checkerboard-based [8] and motion-based [39] extrinsic calibration approaches. It is noteworthy that most existing single-point LiDAR calibration methods assume the laser pointer is always visible [25], [22]. They also rely heavily on either accurate odometry [27] or structured pattern detection [23], requiring a large number of images for extrinsic calibration. We cannot apply these approaches because the proposed method targets an invisible laser dot and

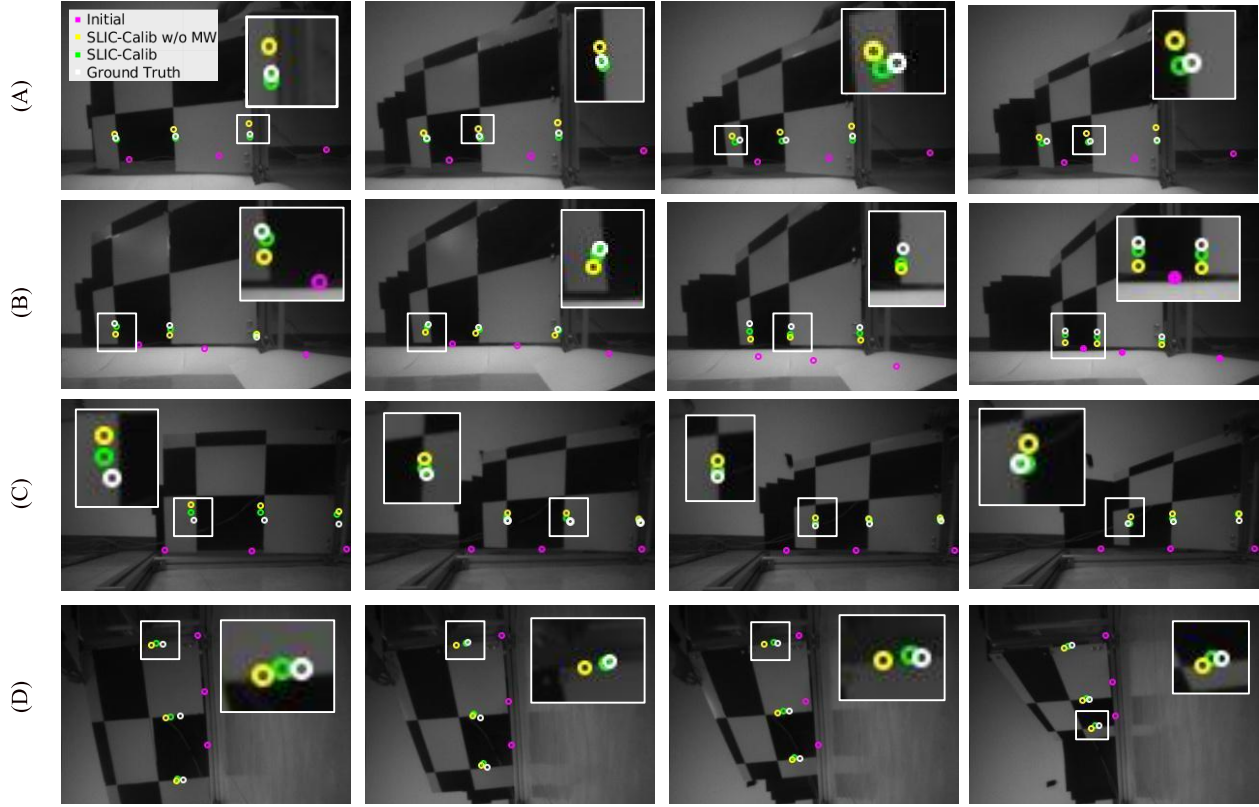


Fig. 10. We visualize the reprojection on the image plane of three points from the single-point LiDAR based on the estimated extrinsic parameters. For each configuration, the target board gradually opens over time from the left to the right column. Different colored points visualize the reprojection point: magenta points represent the result using initial extrinsic parameters, yellow points represent results from point-line constraints optimization, green points represent results from MW refinement, and white points serve as the ground-truth.

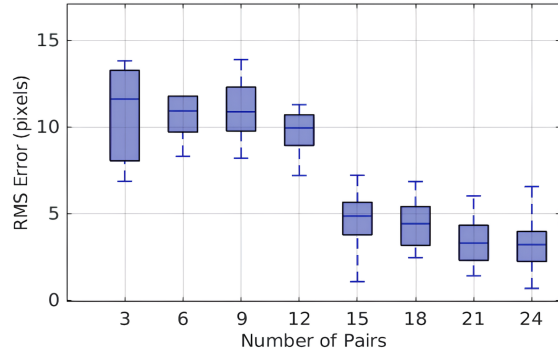


Fig. 11. Accuracy of reprojection error of the proposed method versus the number of LiDAR and camera pairs. Extrinsic calibration stabilizes with about 15 or more LiDAR-camera pairs.

operates in minimally textured environments with extremely sparse calibration data.

MiniVO [8] relies on checkerboard images and requires at least 40 pairs for calibration, making data collection cumbersome and impractical for lightweight applications due to the difficulty of obtaining sufficient calibration data. The board must be positioned at various angles while staying within the single-point LiDAR’s measurement range, making it difficult to acquire enough valid calibration pairs. Table III shows the results of the proposed SPLiCE and MiniVO [8]. The smallest error for each dataset is bolded. Our method

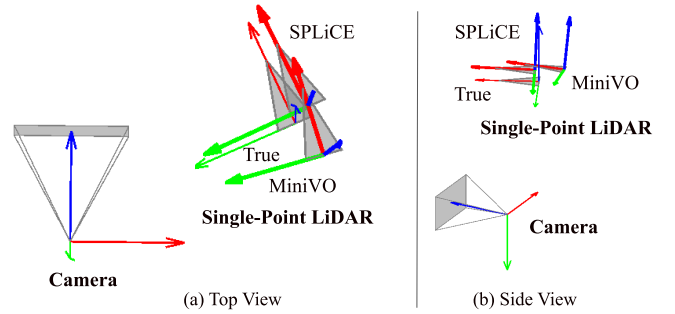


Fig. 12. Qualitative comparison results of the proposed SPLiCE, MiniVO [8], and the ground truth LiDAR poses with respect to the camera frame. The proposed method shows more qualitatively accurate extrinsic calibration results.

achieves higher accuracy with only about 15 valid pairs, whereas MiniVO requires more than three times as many in a controlled environment. Fig. 12 also shows this improvement, demonstrating the superior alignment of our estimated extrinsic parameters with the ground truth from OptiTrack.

To implement [39], we mount the flow deck to the Crazyflie and run visual-inertial odometry (VIO) to simulate a virtual 2D LiDAR sensor like Hokuyo UST-10LX by accumulating single-point LiDAR data along estimated trajectories from VIO. It calibrates the sensor by minimizing reprojection error, associating planes in the camera frame with laser points on the intersection line through a linear least-

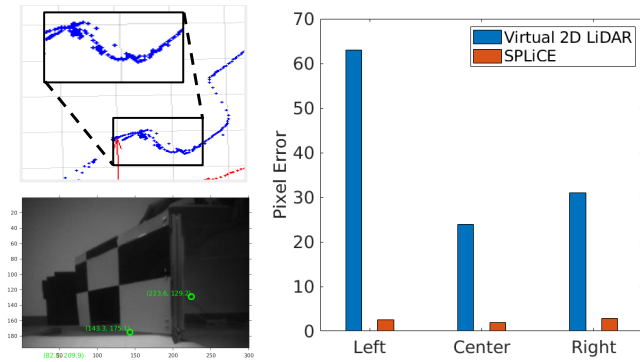


Fig. 13. The blue dots represent the point clouds accumulated by the single-point LiDAR data, and the red dots denote the estimated Crazyflie poses from VIO. Green dots represent the inaccurate reprojection results using extrinsic parameters from the virtual 2D LiDAR approach [39].

squares problem. Fig. 13 shows the comparison results with the proposed SPLiCE. The extrinsic parameters estimated from virtual 2D LiDAR calibration result in an average pixel error of 42 pixels, while our SPLiCE reduces it to 3 pixels. Small drift errors from VIO contaminate the accumulated point cloud, misaligning 3D-2D correspondences and leading to severe inaccurate extrinsic calibration.

VI. CONCLUSION

We present a novel calibration method for single-point LiDAR and a camera calibration algorithm, introducing a specially designed rotating calibration board. The target board contains holes that accumulate single-direction LiDAR data to represent the MW frame, enabling estimation of three points. Although the camera does not recognize the holes in the images, the black-and-white grid pattern allows the detection of only the lines on the target board. After determining the extrinsic parameters using point-line and point-point distances, we improve accuracy by utilizing the regularities of the MW structure. We evaluate the calibration for intra-drone and inter-drone cases and compare the performance using OptiTrack. Future works involve integrating the proposed calibration into SLAM frameworks for onboard use in micro drones, enabling localization and mapping with sparse or single-direction LiDAR.

REFERENCES

- [1] Z. Pusztai and L. Hajder, "Accurate calibration of lidar-camera systems using ordinary boxes," in *IEEE ICCV-W*, 2017.
- [2] L. Zhou, Z. Li, and M. Kaess, "Automatic extrinsic calibration of a camera and a 3d lidar using line and plane correspondences," in *IEEE IROS*, 2018.
- [3] X. Lv, B. Wang, Z. Dou, and S. Wang, "Lccnet: Lidar and camera self-calibration using cost volume network," in *IEEE CVPR*, 2021.
- [4] V. Niculescu, T. Polonelli, M. Magno, and L. Benini, "Nanoslam: Enabling fully onboard slam for tiny robots," *IEEE IoT*, 2023.
- [5] C. Friess, V. Niculescu, T. Polonelli, M. Magno, and L. Benini, "Fully onboard slam for distributed mapping with a swarm of nano-drones," *IEEE IoT*, 2024.
- [6] J. Wither, C. Coffin, and T. Hollerer, "Fast annotation and modeling with a single-point laser range finder," in *IEEE ISMAR*, 2008.
- [7] R. Giubilato, S. Chiodini, and S. Debei, "Scale correct monocular visual odometry using a lidar altimeter," in *IEEE IROS*, 2018.
- [8] R. Giubilato, S. Chiodini, M. Pertile, and S. Debei, "Minivo: Minimalistic range enhanced monocular system for scale correct pose estimation," *IEEE Sensors*, 2020.

- [9] H. Xu, X. Liu, and Y. Liu, "A novel visual scale recovery for resource-limited platform using single-point laser and spatial constraints," *IEEE Sensors*, 2023.
- [10] Y. Xu, X. Zhu, J. Shi, H. Bao, and H. Li, "Depth completion from sparse lidar data with depth-normal constraints," in *IEEE ICCV*, 2019.
- [11] J. Park, Y.-J. Li, and K. Kitani, "Flexible depth completion for sparse and varying point densities," in *IEEE CVPR*, 2024.
- [12] W. Giernacki, M. Skwierczyński, W. Witwicki, P. Wroński, and P. Kozierski, "Crazyflie 2.0 quadrotor as a platform for research and education in robotics and control engineering," in *IEEE MMAR*, 2017.
- [13] V. Fremont, P. Bonnifait, *et al.*, "Extrinsic calibration between a multi-layer lidar and a camera," in *IEEE MFI*, 2008.
- [14] H.-J. Chien, R. Klette, and U. Franke, "Visual odometry driven online calibration for monocular lidar-camera systems," in *IEEE ICPR*, 2016.
- [15] J. Levinson and S. Thrun, "Automatic online calibration of cameras and lasers," in *RSS*, 2013.
- [16] S. Mishra, G. Pandey, and S. Saripalli, "Extrinsic calibration of a 3d-lidar and a camera," in *IEEE IV*, 2020.
- [17] Y. Zhu, C. Li, and Y. Zhang, "Online camera-lidar calibration with sensor semantic information," in *IEEE ICRA*, 2020.
- [18] Z. Luo, G. Yan, and Y. Li, "Calib-anything: Zero-training lidar-camera extrinsic calibration method using segment anything," *ArXiv*, 2023.
- [19] N. Schneider, F. Piewak, and U. Franke, "Regnet: Multimodal sensor registration using deep neural networks," in *IEEE IV*, 2017.
- [20] K. Yuan, Z. Guo, and Z. J. Wang, "Rgnet: Tolerance aware lidar-camera online calibration with geometric deep learning and generative model," *IEEE RA-L*, 2020.
- [21] Y. Xiao, Y. Li, and Y. Zhang, "Calibformer: A transformer-based automatic lidar-camera calibration network," in *IEEE ICRA*, 2024.
- [22] T. Nguyen and G. Reitmayr, "Calibrating setups with a single-point laser range finder and a camera," in *IEEE IROS*, 2013.
- [23] F. Wang, H. Dong, Y. Chen, and N. Zheng, "An accurate non-cooperative method for measuring textureless spherical target based on calibrated lasers," *Sensors*, 2016.
- [24] Z. Liu, D. Lu, W. Qian, G. Gu, J. Zhang, and X. Kong, "Extrinsic calibration of a single-point laser rangefinder and single camera," *Optical and Quantum Electronics*, 2019.
- [25] J. Kang, B. Wu, and Y. Wang, "Calibration method of 1d laser rangefinder for large-scale 3d measurement under harsh environment," *IEICE Electronics Express*, 2020.
- [26] Y. Tu, Y. Song, F. Liu, Y. Zhou, T. Li, S. Zhi, and Y. Wang, "An accurate and stable extrinsic calibration for a camera and a 1d laser range finder," *IEEE Sensors*, 2022.
- [27] R. Ishikawa, S. Zhou, Y. Sato, T. Oishi, and K. Ikeuchi, "Lidar-camera calibration using intensity variance cost," in *IEEE ICRA*, 2024.
- [28] C. Yuan, X. Liu, X. Hong, and F. Zhang, "Pixel-level extrinsic self calibration of high resolution lidar and camera in targetless environments," *IEEE RA-L*, 2021.
- [29] J. Zhang, Y. Liu, M. Wen, H. Zhang, and D. Wang, "L2v2t2calib: Automatic and unified extrinsic calibration toolbox for different 3d lidar, visual camera and thermal camera," in *IEEE IV*, 2023.
- [30] Z. Huang, X. Zhang, A. Garcia, and X. Huang, "A novel, efficient and accurate method for lidar camera calibration," in *IEEE ICRA*, 2024.
- [31] J. Beltr'an and F. A. Garc'ia, "Automatic extrinsic calibration method for lidar and camera sensor setups," *IEEE T-ITS*, 2021.
- [32] G. Yan, F. He, C. Shi, P. Wei, X. Cai, and Y. Li, "Joint camera intrinsic and lidar-camera extrinsic calibration," in *IEEE ICRA*, 2023.
- [33] D. P. Paudel and P. Vasseur, "2d-3d synchronous/asynchronous camera fusion for visual odometry," *AURO*, 2019.
- [34] J. M. Coughlan and A. L. Yuille, "Manhattan world: Compass direction from a single image by bayesian inference," in *IEEE ICCV*, 1999.
- [35] P. Kim, H. Li, and K. Joo, "Quasi-globally optimal and real-time visual compass in manhattan structured environments," in *IEEE RA-L*, 2022.
- [36] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "Lsd: A fast line segment detector with a false detection control," *IEEE T-PAMI*, 2010.
- [37] J.-C. Bazin and M. Pollefeys, "3-line ransac for orthogonal vanishing point detection," in *IEEE IROS*, 2012.
- [38] S. Sim, J. Sock, and K. H. Kwak, "Indirect correspondence-based robust extrinsic calibration of lidar and camera," *Sensors*, 2016.
- [39] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in *IEEE IROS*, 2004.