Image and Vision Computing 155 (2025) 105394

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Accurate LiDAR–camera calibration using feature edges

YunFeng Hua [a,d], QinYu Liu [c,d], TengFei Jiang [d], Jian Zhang [d], WeiWei Xu [a,*], Yan Tian [b,*]

[a] College of Computer Science and Technology, Zhejiang University, HangZhou, 310058, China
[b] School of Computer Science and Technology, Zhejiang Gongshang University, HangZhou, 310018, China
[c] School of Mechatronics Engineering, Harbin Institute of Technology, Heilongjiang, 150001, China
[d] SHINING 3D TECH CO., LTD., HangZhou, 311258, China

ARTICLE INFO

Keywords:
Calibration
SLAM
Multi-sensor
Computer vision

ABSTRACT

Accurate extrinsic calibration is essential for minimizing alignment errors between LiDAR and cameras, ensuring precise sensor data registration, and enhancing the robustness and accuracy of Simultaneous Localization and Mapping (SLAM) systems. Although previous calibration techniques employing plane and point features, as well as neural networks, have shown promise, they are not devoid of limitations. Particularly, prominent point features such as chessboard corners may lack precise counterparts in the sparse point clouds generated by LiDAR. To address these challenges, we introduce a novel LiDAR–camera extrinsic calibration method that leverages edge registration with known correspondences. This approach significantly reduces calibration discrepancies associated with imprecise correspondences and systemic noise, providing a systematic and rigorous framework to improve the precision of extrinsic calibration between LiDAR and camera systems. Additionally, the flexibility of our method allows for the use of common everyday objects, such as boxes, books, or sheets of paper, for calibration purposes, simplifying the procedure and enhancing its practical applicability.

1. Introduction

Multi-sensor systems that integrate LiDAR and camera technologies are essential for enhancing three-dimensional perception and environmental understanding. LiDAR is capable of acquiring high-precision 3D information, exhibiting strong resistance to interference from lighting and weather conditions, yet it is limited by its low resolution and lack of color information. On the other hand, cameras provide high-resolution color information but are more susceptible to interference and variations in lighting. The complementary nature of LiDAR and camera data fusion allows for the acquisition of richer information. Compared to single-sensor systems, these integrated systems offer deeper environmental insights, enabling a suite of critical applications, including spatial localization [1,2], obstacle detection and tracking [3–6], map building and scene reconstruction [7,8], object recognition and classification [9,10], and pose estimation and motion analysis [11–13]. The prerequisite for multi-sensor fusion is the accurate determination of extrinsic parameters, specifically the rotation matrix and translation vector between the LiDAR and camera coordinate systems. However, precisely calibrating multiple sensors within a system remains a significant challenge.

In the external calibration of multi-sensor systems, establishing accurate correspondences between features across different sensor modalities is crucial. Despite the existence of various methods proposed in the literature [14–22], the accuracy of point feature and plane feature matching methods remain challenges in the field. For instance, in sparse point clouds captured by LiDAR, it may be difficult to find precise correspondences for salient point features such as chessboard corners. On the other hand, plane features [23,24] are relatively easier to establish correspondences, but their constraints are not strong enough, leading to less accurate calibration results. In contrast, edge features not only provide rich structural information but also exhibit high robustness to illumination changes and noise, combining the advantages of both point and plane features. Therefore, the introduction of edge features offers a more accurate and robust solution for the calibration of multi-sensor systems, facilitating higher-quality 3D environment perception and understanding.

To this end, we present a novel calibration method that uses edge registration with known correspondences to refine the extrinsic parameters between the LiDAR and camera systems. By concurrently capturing data from LiDAR and camera sensors, we elaborately extract the edges of target objects from both datasets. For image edges, to expedite the identification of target locations, we initially adopt a manual annotation approach to obtain a rough outline, which is then refined by fitting it with the edge contours extracted from the image, thereby achieving precise extraction. As for the point cloud edges, considering that the calibration board can be perceived as two-dimensional, we

* Corresponding author.
E-mail addresses: xww@cad.zju.edu.cn (W. Xu), tianyan@zjgsu.edu.cn (Y. Tian).

https://doi.org/10.1016/j.imavis.2024.105394
Received 24 April 2024; Received in revised form 17 September 2024; Accepted 11 December 2024
Available online 25 January 2025

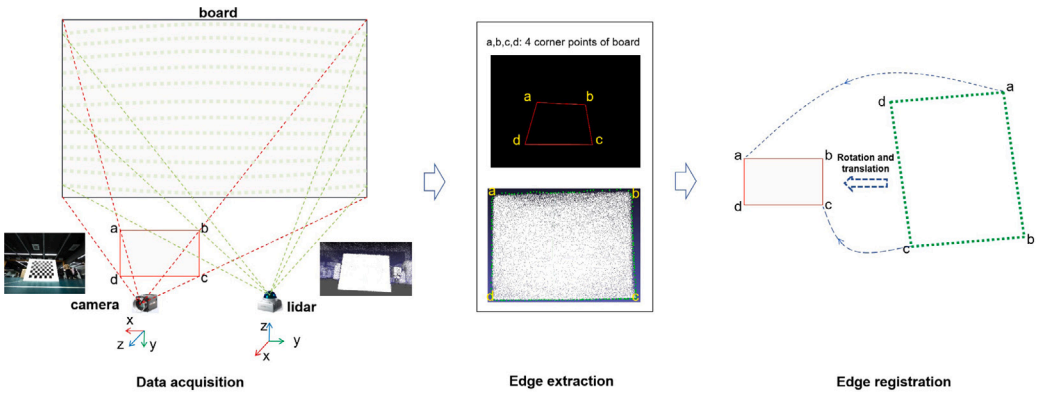

Fig. 1. An illustrative diagram of the overall calibration process, from data acquisition to obtaining the extrinsic parameters through edge registration.

can directly apply a convex hull algorithm to acquire the edge contour of the target object. Moreover, given the significant noise inherent in the point cloud, we further enhance the edge extraction by fitting it with the ground truth contour derived from the calibration board. Utilizing edge registration process, we are able to compute the extrinsic parameters that articulate the spatial relationship between the sensors.

The framework of our proposed system is depicted in Fig. 1, offering a visual overview of the system's architecture and workflow. Our key contributions include:

- We have established a complete calibration system capable of efficiently extracting the edges of the target calibration board from both image and point cloud.
- By optimizing poses using corresponding edges, we can achieve a more accurate determination of the rigid transformation between sensors.
- Calibration can be efficiently executed utilizing a standard calibration board, or alternatively, with the simplicity of everyday objects such as a box or a book.

2. Related work

In the domain of multi-sensor system calibration, an array of innovative methodologies has been proposed, each making a significant contribution to the advancement of calibration techniques [25–29]. The pursuit of enhanced precision and robustness in calibration has led to the recent emergence of numerous approaches, particularly targeting the calibration of LiDAR and camera systems.

Feature-Based Calibration One prominent approach in this field capitalizes on feature extraction and matching techniques, focusing on planar-based or point-based features. Notably, Zhang et al. [14] and Pandey et al. [15] have advocated for the use of planar feature constraints to refine the extrinsic parameters of LiDAR–camera systems. Their methodology involves fitting the corners detected by the camera to a plane and aligning the corresponding 3D points with this plane, effectively using geometric scene constraints to bolster calibration accuracy. Conversely, Huang et al. [17] have explored a different avenue, employing targets with known dimensions and geometries to enhance the estimation of target pose. By leveraging the targets' known dimensions, they address the quantization and systematic errors present in LiDAR data, thereby improving the precision of the calibration process. This technique shows particular advantage in scenarios characterized by targets with distinct shapes and sizes. Beltran et al. [19] proposed to extract reference points from additional sensor data for accurate calibration. By registering these points, they determine an optimal rigid transformation, demonstrating the efficacy of their approach in managing sensors of varying resolutions and orientations. They have engineered a calibration board that captures both two-dimensional and three-dimensional features, facilitating calibration through the use of pre-established correspondences.

Reflectivity-Based Calibration Furthermore, the reflectivity intensity of LiDAR point clouds has been recognized for its utility in corner point extraction [16,18]. The variability in point reflectivity, influenced by factors such as material properties, incident angles, and colors, provides valuable information for feature extraction. Wang et al. [16] introduced a fully automated and accurate extrinsic calibration method for 3D LiDAR and camera systems based on laser reflection intensity. Cui et al. [18] present a fully automated calibration technique for non-repetitive scanning Solid-State LiDAR (SSL) and camera systems. A temporal-spatial-based geometric feature refinement, and reflectance intensity distribution-based 3D corner estimation pipeline for point clouds from non-repetitive scanning SSL is presented. Despite the promise of this method, the sparsity of point cloud data presents a significant challenge in the precise detection of corners, which is crucial for the reliability of extrinsic calibration. Our method leverages edges with known correspondences for calibration, mitigates the impact of point cloud sparsity.

Deep Learning-Based Calibration Beyond traditional techniques, there is a growing interest in applying deep learning to sensor calibration tasks [30–33]. Schneider et al. [30] present RegNet, which casts all three conventional calibration steps (feature extraction, feature matching and global regression) into a single real-time capable CNN. Shi et al. [31] present CalibRCNN, not only uses the LSTM network to extract the temporal features between 3D point clouds and RGB images of consecutive frames, but also uses the geometric loss and photometric loss obtained by the interframe constraint to refine the calibration accuracy of the predicted transformation parameters. And Lv et al. [33] propose an online LiDAR–camera Self-calibration Network (LCCNet), different from the previous CNN-based methods. LCCNet can be trained end-to-end and predict the extrinsic parameters in real-time. Deep neural networks, which have shown success across various computer vision tasks, may offer a new avenue for addressing the extrinsic calibration challenge. Training a neural network on a comprehensive dataset of calibrated sensor pairs could potentially enable the learning of complex mappings between sensor inputs and the desired extrinsic parameters. This approach could simplify the calibration process and enhance its accuracy, marking a significant step forward in the field. Nonetheless, the refinement of calibration accuracy using deep learning methods remains an area ripe for further exploration.

In light of the aforementioned methods, it is clear that the extraction of point cloud features, is a pivotal component of extrinsic calibration, directly affecting the calibration outcome. Building upon this understanding, this paper introduces a novel extrinsic calibration method for LiDAR and camera systems. Our approach is based on known corresponding edges, leveraging the extraction of corresponding edge features from both LiDAR and camera data to optimize extrinsic parameters. This method not only enhances the precision of extrinsic calibration but also addresses the challenge of point cloud sparsity, offering a streamlined and effective solution to the calibration process.

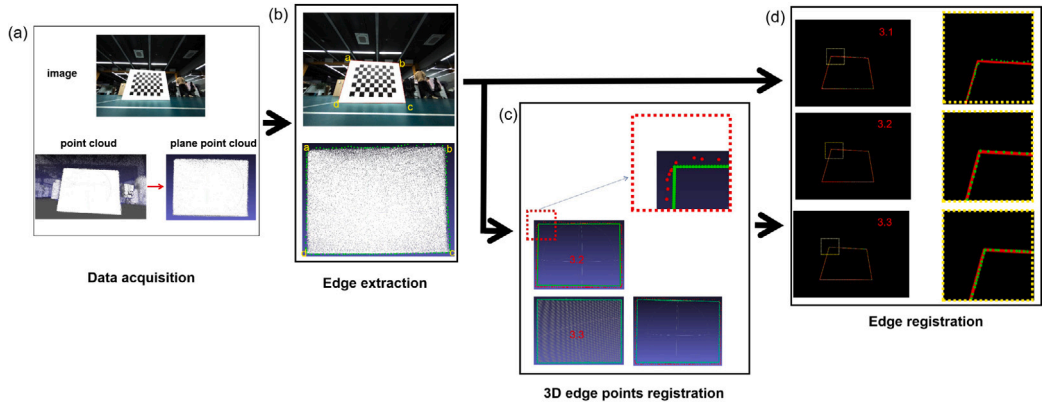

Fig. 2. (a) Input data from camera and LiDAR, the plane point cloud is extracted from the overall point cloud. (b) Edge extraction. (c) 3D edge points registration. Red points: edge extraction from plane point cloud. Green points: edge extraction from ground-truth calibration board. (d) Edge registration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Our approach

Our approach involves optimizing the extrinsic parameters [**R T**] between the LiDAR and camera by identifying corresponding 3D–2D edges. For this, we employ an iterative optimization method to refine the 3D points from LiDAR point cloud and 2D edge lines from RGB image. The objective function is defined as follows:

$$\underset{\mathbf{R}\in SO(3),\mathbf{T}\in\mathbf{R}^3}{\operatorname{argmin}} \sum_{\mathbf{L}_j, X_{edge}} \|e(\mathbf{L}_j, \mathbf{R}X_{edge} + \mathbf{T})\|^2, \tag{1}$$

where **R** and **T** represents rotation and translation from LiDAR to camera. **P** is the projection operation from 3D to 2D on the image plane and X_{edge} are points sampled from the edges of plane point cloud. $\mathbf{L}_j$ is the corresponding edge line in RGB image.

Plücker coordinates. An edge line can be represented by the Plücker coordinates [34]. Given two points $\mathbf{p}_{l1}$ and $\mathbf{p}_{l2}$ on an edge line, the Plücker coordinates of the edge line have the form:

$$\mathbf{L} = [\mathbf{m}; \mathbf{d}], \mathbf{m} = \mathbf{p}_{l1} \times \mathbf{d} \ \textit{and} \ \mathbf{d} = \frac{\mathbf{p}_{l2} - \mathbf{p}_{l1}}{\|\mathbf{p}_{l2} - \mathbf{p}_{l1}\|_2} \tag{2}$$

where $\times$ represents the cross product. The Plücker coordinates are homogeneous coordinates. Here we normalize d, as this can lead to more concise formulas in the following description. Using the Plücker coordinates in (2), we can write the distance from a point p to L as

$$e(\mathbf{L}, \mathbf{p}) = \mathbf{m} - \mathbf{p} \times \mathbf{d} \tag{3}$$

To enhance the optimization process, we propose a method that aligns edge points from the fused LiDAR point cloud with image edges. However, the presence of noise in the LiDAR data presents a challenge as it can introduce errors during the edge extraction process. To overcome this issue, we utilize two types of measurement tools to obtain accurate edge points, which serve as the ground truth for the calibration board. The first tool is a standard tape measure used to measure the length of the board's edges, enabling us to construct a quadrilateral with precise dimensions. The second tool is a high-precision scanner capable of capturing 3D data with an accuracy of less than 0.02 mm. These meticulously collected ground truth data are registered into the LiDAR point cloud's coordinate system and subsequently employed in the image edge registration process, thereby improving the accuracy of the calibration results. In the following sections, we will discuss the calibration process using the following methodologies: (i) direct calibration using derived edge points from the extracted plane point cloud, (ii) calibration using points constructed based on tape measure-based edge length measurements, and (iii) calibration using edge points obtained from the scanner as the ground truth of the calibration board.

3.1. Calibration with a unknown-sized board

In a controlled experimental scenario, a calibration board is positioned to ensure that both the LiDAR and camera systems capture the board within their respective fields of view simultaneously. Our system framework, illustrated in Fig. 1, includes three main modules: (1) Data acquisition. (2) Edge extraction. (3) Edge registration.

Data acquisition. In the deployment of the LiDAR–camera system within the calibration scene, the equipment is initiated in a stationary state to simultaneously acquire image and point cloud data. Throughout the calibration process, the relative positioning between LiDAR and camera is consistently maintained to ensure a stable reference frame. Given the intrinsic sparsity characteristic of LiDAR point clouds, achieving comprehensive data capture of the calibration target necessitates the deliberate movement of the LiDAR–camera system, thereby facilitating the acquisition of a complete and dense point cloud representation of the calibration board.

Subsequently, the initial frame of point cloud data serves as the world coordinate, with all subsequent point clouds being registered to this established coordinate system. This crucial registration process is facilitated through the utilization of the widely employed Iterative Closest Point (ICP) algorithm, which iteratively refines the transformation by minimizing the distance between corresponding points in the two point clouds, thereby ensuring precise alignment.

Moreover, the integration of Inertial Measurement Unit (IMU) data has been identified as a key mechanism for enhancing motion estimation and augmenting the efficacy of point cloud alignment [35, 36]. The fusion of IMU sensor data with LiDAR inputs enables a more accurate estimation of the sensor's motion, consequently yielding improved alignment results. During the process, if the movement is sufficiently slow, the incorporation of IMU data has minimal impact on the registration results. The primary role of the IMU is to compensate for motion-induced point cloud offsets, ensuring accurate alignment. Fig. 2(a) visually depicts an image captured by the camera alongside an integrated plane point cloud subsequent to the registration process.

Edge extraction. Following the acquisition of the scene data, the process of edge extraction is initiated independently for both the image and point cloud.

In the context of calibration scenarios, the selection of image edge extraction techniques is contingent upon the specific characteristics of the background. In environments with a distinct and homogeneous backdrop, such as a green screen or a monochromatic wall, the Canny edge detector can be effectively utilized for edge detection, succeeded by contour detection algorithms to outline the perimeter of the calibration board. This approach leverages the simplicity of the background

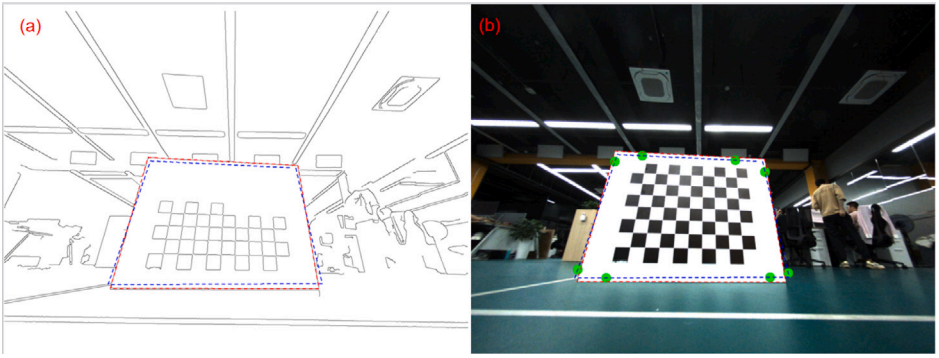

Fig. 3. Edge extraction: (a) the edge contours of the image, (b) the green points are manually marked edge points, the blue lines represent the edge contours formed by connecting the edge points, and the red lines indicate the final optimized edges after fitting. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to enhance the accuracy of edge delineation.

Conversely, in the absence of a suitable background, the initial contour edge is delineated through manual intervention, serving as the foundation for subsequent edge refinement. This refinement process is elaborately executed to ensure that the extracted edges are precisely aligned with the physical boundaries of the calibration board. As illustrated in Fig. 3, the edge extraction process is exemplified within a complex background scenario. Fig. 3(a) presents the contours extracted directly from the image, revealing that while most of the calibration board's edge contours have been identified, some edges are connected to the background due to its complexity. Simple manual intervention quickly distinguishes the overall edge contour lines. Fig. 3(b) highlights the manually annotated edge points in green. The blue lines represent the preliminary edge contours constructed by connecting these annotated points. The red lines denote the refined edges post-optimization, which are the outcome of a fitting procedure aimed at enhancing the accuracy of edge representation.

In this process, four edge lines are optimized, each line L is represented as Plücker coordinates [**m**;**d**]. Specifically, within Fig. 3(b), fifty points are uniformly sampled along each of the blue lines and paired with the nearest points from Fig. 3(a) to form the target point set $\boldsymbol{p}_t$. This set serves as the basis for optimizing the values of [m;d]. The objective function is outlined as follows:

$$\underset{\mathbf{m},\mathbf{d}\in\mathbf{R}^3}{\operatorname{argmin}} \sum_{\mathbf{p}_t} \|e(\mathbf{L},\mathbf{p}_t)\|^2, \tag{4}$$

where **m**, **d** represent the Plücker coordinates of the edge line.

For the precise extraction of edges from the point cloud, it is imperative to eliminate all points that do not correspond to the target plane across the entire data set. Fig. 4 provides a visual representation of the results from the point cloud extraction of the calibration board. In Fig. 4(a), the comprehensive point cloud of the scene is depicted, showcasing the full range of captured data. Fig. 4(b) presents the point cloud specifically associated with the calibration board, highlighted in red. This subset has been filtered based on the analysis of normal vectors and spatial coordinates. The initial step in the process involves the identification of all planar structures within the point cloud, followed by the removal of non-continuous points to preserve the most extensive connected segment. Continuing with the refinement, the filtered plane point clouds are employed to determine the orientation and position of the plane. Subsequently, constraints based on distance and normal vector analysis are applied to precisely locate the target plane. Ultimately, the filtered point cloud, treated as effectively two-dimensional, is processed using a concave hull detection algorithm to extract the edge points. The edges extracted from the planar point cloud are visually depicted in Fig. 2(b).

Edge registration. The calculation of the extrinsic parameters [**R T**] is predicated on identifying multiple sets of corresponding 3D edge points from the LiDAR point cloud and their 2D counterparts from the camera image. These correspondences facilitate the determination of the extrinsic parameters using Eq. (1). Initially, an estimation of the extrinsic parameters is established, which is refined iteratively to improve precision. This refinement is guided by the alignment of the edge points, ensuring that the computed parameters more closely match the actual spatial relationship between the LiDAR and the camera.

The initial step in the extrinsic parameter estimation process involves establishing a correspondence between the four corner points of a plane point cloud and their associated edge lines in the image. Each corner point can be related to two distinct edge lines. This correspondence is fundamental for deriving the initial transformation matrix, which serves as a precursor for further refinement of the parameters. If an initial transformation matrix is already established, this step can be omitted, allowing for a direct transition to the subsequent phase of the procedure.

The second step entails projecting all points of the plane point cloud edge onto the UV coordinate system based on the initial transformation matrix as follows:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & \gamma & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{pmatrix} * \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{5}$$

Eq. (5) illustrates the projection of the 3D points after applying the [**R T**] transformation. Then, search for the nearest neighboring edge of the image, and iteratively optimize the rigid transformation [**R T**] by fitting the edge in accordance with Eq. (1).

Fig. 2(d) demonstrates that following 15 iterations, there is a close alignment between the edge points of the planar point cloud and the edges detected in the image.

3.2. Calibration with a known-sized board

In practical applications, the accuracy of calibration can be significantly affected by the presence of noise in LiDAR data. To mitigate this challenge, a robust approach involves acquiring ground-truth data from the calibration target as an initial step. This ground-truth data provides a reliable benchmark for the actual spatial relationship between the LiDAR and the camera, serving as a reference for comparison. Leveraging this ground-truth data enables the refinement and verification of the transformation matrix during optimization, thereby enhancing the accuracy and robustness of the calibration process.

Given the rectangular shape of our calibration board, we can directly construct the rectangle by measuring its edge lengths. Tools such as rulers and tape measures enable the straightforward acquisition of precise measurements, typically at the millimeter level. The

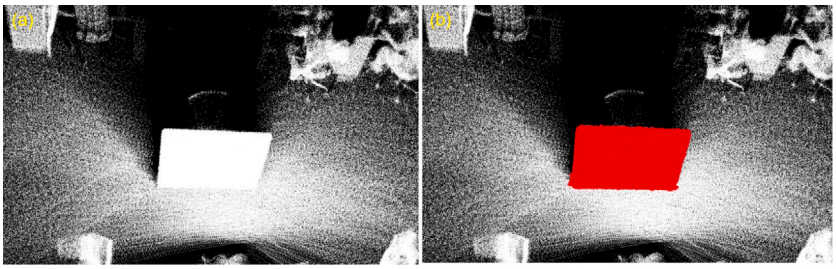

Fig. 4. Plane extraction: (a) the point cloud of the calibrate scene, (b) the red points represent the plane extracted from the scene point cloud. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

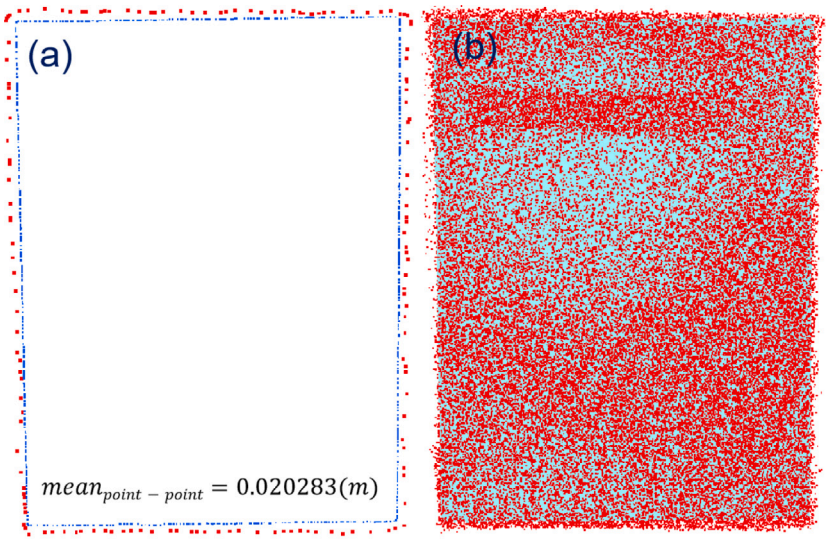

Fig. 5. The ground truth of board scanned by Shining3d FreeScan Trio [https://www.shining3d.cn/gongye3d/shouchi/FreeScanTrio] with an error of only 0.02 mm. The mean registration error with the extracted contour from the LiDAR point cloud is about 2 cm.

accessibility and accuracy of these tools make them well-suited for obtaining reference measurements, effectively overcoming the challenges associated with acquiring ground-truth data for the calibration board.

3D edge registration. Using a standard tape measure, we obtained measurements for the edges of the calibration board with an impressive margin of error of less than 1 millimeter. Utilizing these measurements, we constructed a quadrilateral that serves as a reliable geometric representation of the calibration board. The upper left corner of the board was designated as the reference origin, and the board's edges were aligned with the x and y axes. The complete structure of the board was meticulously assembled based on the precisely measured dimensions. This quadrilateral, constructed using measured dimensions, exhibited superior accuracy compared to the edge points directly extracted from the plane point cloud.

Subsequently, we needed to transform the constructed quadrilateral to the coordinate system of the plane point cloud. By identifying the four corners of these two sets of edge points as corresponding points and calculating an initial transformation matrix, we could effectively initiate the optimization process with the following steps:

$$\underset{\mathbf{R}_{MP}\in SO(3),\mathbf{T}_{MP}\in\mathbf{R}^3}{\operatorname{argmin}} \quad \|\mathbf{R}_{MP}X_M + \mathbf{T}_{MP} - X_P\|^2 \tag{6}$$

where X_p are edge points of plane point cloud and X_M represent tape measured-edge-points. $\mathbf{R}_{MP}$ and $\mathbf{T}_{MP}$ is rotation and translation transformation from tape measured-edge-points coordinate system to plane point cloud's. Fig. 2(c) illustrates the fitting outcome, the red points correspond to the edge points extracted from the plane point cloud via edge extraction, while the green points indicate the tape measured-edge-points. As depicted in Fig. 2(c), the data procured from the LiDAR scan exhibits significant overflow at the edges, with the extracted edge data appearing considerably larger than the actual dimensions. Consequently, direct calibration utilizing the scan data as described in Section 3.1 would result in substantial errors.

Following the registration of the tape measured-edge-points to the coordinate system of the plane point cloud, a 2D Edge registration process, as described in Section 3.1, is implemented to derive the rigid transformation matrix [**R T**]. As illustrated in Fig. 2(d), the outcome of the 2D Edge registration process is visually apparent, showcasing a closer alignment between the measured edge points and the fitted edges within the image space.

3.3. Calibration with a known-ground-truth board

With the continuous advancements in scanning technology, the attainment of highly precise calibration results has become increasingly feasible. In Fig. 5, we present the ground-truth data of the calibration board, which was scanned using the state-of-the-art Shining3d FreeScan Trio. This scanning device has demonstrated an exceptional accuracy level, achieving measurements with an impressive precision of up to 0.02 mm. After the registration of the ground truth of the calibration board with the extracted contour from the LiDAR point cloud, as shown in Fig. 5, the average point-to-point distance is approximately 2 cm. Leveraging this captured scanned data as our ground truth, we embarked on a comprehensive refinement process to further enhance the calibration results, aiming to achieve optimal performance in our system.

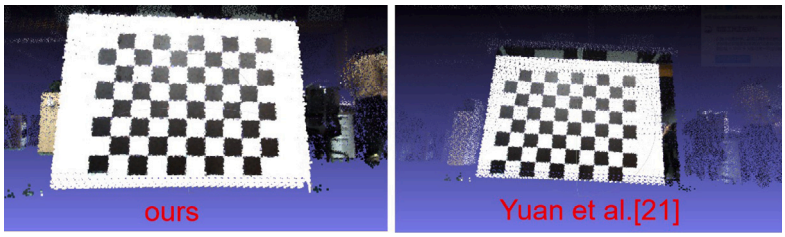

Fig. 6. Comparison between ours and Yuan et al. [37].

Table 1
Rotation error with various noise intensity. Unit (rad).

Gaussian noise (m)	Zhou et al. [23]	Singandhupe et al. [24]	Ours(unknown size)	Ours(known size)
0.005	0.02431	0.03635	0.01554	**0.00551**
0.01	0.02815	0.03967	0.01893	**0.00559**
0.015	0.03552	0.04029	0.02143	**0.00603**
0.02	0.05857	0.04243	0.04593	**0.01593**
0.025	0.06563	0.04576	0.04723	**0.01876**

Ground-truth fitting. After obtaining the ground truth data of the calibration board, we utilized an Edge extraction technique to outline the edges of the actual calibration board. Subsequently, we followed a methodological approach in accordance with the established 3D Edge registration protocol. This technique enabled the calculation of a transformation matrix with enhanced accuracy, thereby improving the refinement of our calibration model. The results of the fitting process, which encapsulate the synthesis of these intricate methods, are clearly depicted in Fig. 2(d).

4. Experimental results

To validate the effectiveness of our approach, we set up a LiDAR–camera system consisting of a mid-360 LiDAR and an MV-CS016-10UC camera with a resolution of 1440 × 1080 pixels. The intrinsic parameters of the LiDAR are pre-calibrated by the manufacturers, while the intrinsic matrix K of the camera was calibrated using classical methods with chessboard patterns. The calibration board was positioned within the field of view (FOV) of the camera.

For the extraction of contours from the acquired images, we employed the Canny edge detection algorithm, supplemented by the findContours function from the OpenCV library, a conventional and effective approach for contour identification within imagery. In the process of extracting planar point clouds from the comprehensive point cloud data, we initially utilized the SACSegmentation function to filter all planar point clouds, followed by the application of the EuclideanClusterExtraction function to select the largest cluster within the planar point clouds. Based on these points, we calculated the normal and centroid of each plane. Subsequently, we applied distance and normal vector constraints to identify the target plane. Specifically, we set a distance threshold of 2.5 meters and a normal angle constraint of 45 degrees to ensure the target plane was accurately filtered. To further refine the edge extraction from the planar point clouds, we then employed the ConcaveHull algorithm from the Point Cloud Library (PCL). This algorithm excels at identifying the boundary edges of a plane within a point cloud, offering a dependable method for edge detection on a two-dimensional plane.

To evaluate the robustness of our calibration approach, we conducted experiments using both simulated and real data, allowing us to analyze the impact of different levels of noise on the calibration results.

4.1. Simulated data

Using the Maya software, a simulated dataset comprising LiDAR point cloud and image data with predetermined extrinsic parameters between the LiDAR and the camera was generated. Following the introduction of Gaussian noise at various intensities to the point cloud, the edge registration processes outlined in sections 3.1, 3.2, Zhou et al. [23] and Singandhupe et al. [24] were executed to calculate the extrinsic parameters. Subsequently, the obtained results were compared with the ground truth extrinsic parameters, and the error outcomes were visually presented in Tables 1 and 2. The evaluation encompassed the separate computation of the rotation error (in radians) and translation error (in meters).

Our analysis indicates that as noise levels increase, the calibration error for both our method (with unknown dimensions) and Zhou et al. [23] exhibits a consistent upward trend, demonstrating a significant sensitivity to variations in noise levels. Specifically, Zhou's rotational error increased from 0.02431 to 0.06563, and the translational error rose from 0.02892 to 0.14331. In contrast, the error for our method (with known dimensions) and Singandhupe et al. [24] was observed to remain relatively stable and minimal across different noise intensities. Our method's rotational error also showed an increasing trend, yet it was the smallest among the compared methods, while the translational error was largely maintained around 0.01. This comparative assessment highlights the robustness and resilience of our method (with known dimensions) when faced with noise-induced perturbations, confirming its effectiveness in delivering accurate and reliable estimation of extrinsic parameters under challenging environmental conditions.

4.2. Real data

To evaluate the precision of our proposed edge registration-based estimation technique, a comparative analysis was performed against the calibration outcomes reported by Yuan et al. [37], as illustrated in Fig. 6. The results distinctly indicate that while Yuan's method experiences notable discrepancies at the periphery, our approach delivers a more precise alignment. This observation effectively highlights the effectiveness of our approach in accurately estimating extrinsic parameters, particularly in maintaining edge accuracy and overall consistency within the calibration process. Additionally, we conducted extrinsic parameter calibration using a comprehensive dataset that included 1 to 7 sets of data, encompassing a variety of calibration vantage points. The calibrated parameters were then utilized to project the boundary points of the planar point cloud onto the images for the assessment of error metrics. Here, the error is defined as the average distance between the closest image edge points and the projected three-dimensional boundary points based on optimized extrinsic parameters. This measure provides a quantifiable evaluation of the alignment accuracy between the point cloud and image data after calibration.

Table 2
Translation error with various noise intensity. Unit (m).

Gaussian noise (m)	Zhou et al. [23]	Singandhupe et al. [24]	Ours(unknown size)	Ours(known size)
0.005	0.02892	0.04552	0.02573	**0.00869**
0.01	0.05281	0.04965	0.04682	**0.00894**
0.015	0.07862	0.05287	0.06597	**0.01024**
0.02	0.13523	0.05332	0.09217	**0.01173**
0.025	0.14331	0.05458	0.11236	**0.01201**

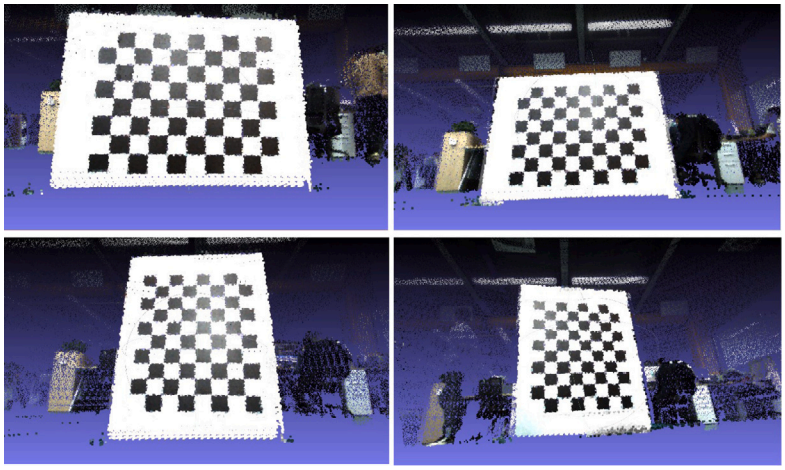

Fig. 7. Point cloud coloring results.

Table 3
Back projection error variation with the number of calibration viewpoints. Unit (pixel).

Views num	Wang et al. [16]	Cui et al. [18]	Zhou et al. [23]	Ours(unknown size)	Ours(known size)
3	3.22177	2.75152	3.12588	2.10215	**1.81054**
4	2.98652	2.10216	2.67095	1.93281	**1.73348**
5	2.55436	1.95445	2.23006	1.91225	**1.60521**
6	2.45158	1.95158	2.18753	1.89274	**1.55367**
7	2.40983	1.90983	2.10863	1.90322	**1.54056**

The calibration outcomes derived from a diverse dataset were utilized for a quantitative analysis of the errors associated with our method under conditions of both unknown and known sizes. As illustrated in Table 3, the error results demonstrate an enhancement in calibration accuracy with an increasing quantity of calibration data points, which substantiates the scalability and efficacy of our approach when applied to a variety of extensive datasets. Furthermore, comparative analyses were conducted with the findings from Wang et al. [16], Cui et al. [18], and Zhou et al. [23]. These comparisons reveal a pattern of consistent improvement in calibration as the range of perspectives widens. Notably, our method's calibration outcomes have surpassed those of the aforementioned studies, indicating that our approach is particularly effective in reducing error and enhancing the precision of extrinsic parameter calibration. In the case of known dimensions, our method's error rates are already among the lowest, and even in situations of unknown dimensions, the error levels are comparatively low. When the number of viewpoints exceeds five, the average error stabilizes at approximately 1.5 pixels. These results reinforce the effectiveness and superiority of our method in the realm of extrinsic calibration and error minimization, highlighting its potential for robust performance across a broad spectrum of applications.

Moreover, we investigated the impact of various point cloud colorization schemes, with the corresponding findings presented in Fig. 7. We also carried out point cloud back-projection onto images, and the results are illustrated in Fig. 8. These visual representations not only serve to validate the accuracy and precision of our proposed approach but also provide a tangible demonstration of its efficacy in facilitating seamless integration between point cloud data and image data.

Overall, our methodology demonstrates superior precision compared to the previously discussed techniques. Specifically, the calibration outcomes of our approach (with known size) exhibit remarkable consistency and are independent of the quantity of calibration viewpoints. Conversely, the convergence of our method (with unknown size) necessitates data acquisition across a diverse array of viewpoints for robust calibration. This refinement ensures that our method is adaptable and reliable across various calibration scenarios.

Additionally, we have tried utilize the intensity information of the point cloud to extract the texture, material, and geometric edges of the scene, and perform fitting calculations with the edges extracted from the RGB image to compute the extrinsic parameters. The results are displayed in the supplementary materials.

5. Conclusion

This paper presents a novel approach to LiDAR–camera extrinsic calibration, utilizing edge registration based on established correspondences to determine extrinsic parameters. This method effectively reduces calibration errors associated with inaccurate correspondences and systemic noise, offering a systematic and rigorous framework to enhance the accuracy of extrinsic calibration between LiDAR and camera systems.

Our method includes a comprehensive error analysis that demonstrates the quantifiable improvements from integrating precise edge points as ground truth. This integration helps overcome the negative impacts of noise, leading to enhanced precision in calibration results. The comparative analysis with existing methods further validates the robustness and effectiveness of our approach, showcasing consistent

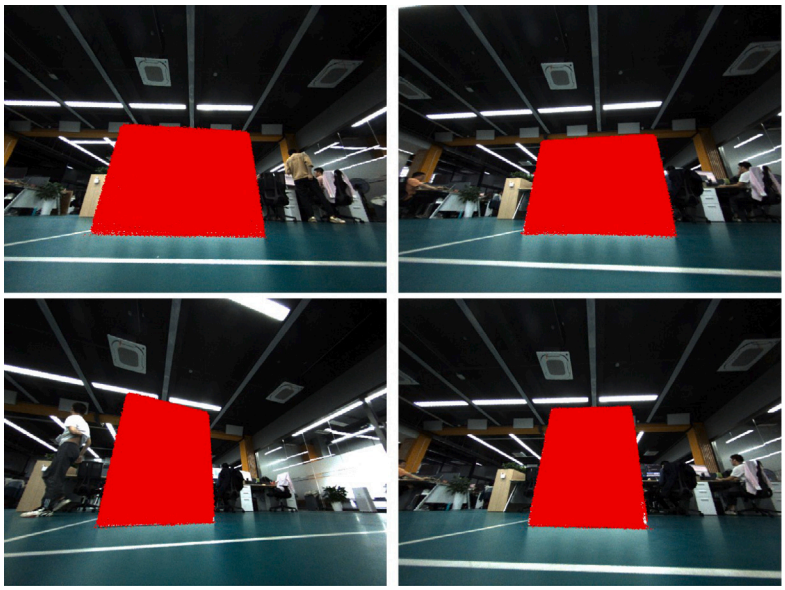

Fig. 8. Point cloud projection results.

improvements with an increasing number of data points and a wide range of perspectives.

In addition to the edge registration technique, we explored the impact of various point cloud colorization schemes and the back-projection of point clouds onto images. These visual validations not only confirm the accuracy and precision of our method but also illustrate its capability to seamlessly integrate point cloud and image data.

The innovative calibration method introduced in this paper is robust and systematic, providing a significant advancement in sensor calibration accuracy, especially in the presence of noise in LiDAR point cloud data. Looking ahead, our future work will explore the use of intensity information within point clouds for edge extraction. Since intensity is influenced by material properties, texture, and structure, edges derived from intensity data are likely to align well with image edges, potentially enabling more effective and automated calibration processes. Furthermore, extending our calibration approach to different types of LiDAR and camera systems will be a focus of our future research.

CRediT authorship contribution statement

YunFeng Hua: Writing – original draft, Methodology, Investigation, Conceptualization. **QinYu Liu:** Methodology, Data curation. **TengFei Jiang:** Writing – review & editing, Methodology, Investigation. **Jian Zhang:** Methodology, Investigation. **WeiWei Xu:** Writing – review & editing, Methodology, Investigation. **Yan Tian:** Writing – review & editing, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Natural Science Foundation of Zhejiang Province (No. Z24F020002), and the Fundamental Research Funds for the Provincial Universities of Zhejiang (No. JRK22003).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.imavis.2024.105394.

Data availability

Data will be made available on request.

References

[1] Z. Chai, Y. Sun, Z. Xiong, A novel method for LiDAR camera calibration by plane fitting, in: IEEE/ASME International Conference on Advanced, 2018, http://dx.doi.org/10.1109/AIM.2018.8452339.
[2] A. Dhall, K. Chelani, V. Radhakrishnan, K. Krishna, Lidar-camera calibration using 3D-3D point correspondences, Comput. Vision Pattern Recognit. (2017) http://dx.doi.org/10.48550/arXiv.1705.09785.
[3] M. Szarvas, U. Sakai, J. Ogata, Real-time pedestrian detection using LIDAR and convolutional neural networks, in: IEEE Intelligent Vehicles Symposium, IV, 2006, pp. 13–15, http://dx.doi.org/10.1109/IVS.2006.1689630.
[4] C. Premebida, U. Nunes, Fusing lidar, camera and semantic information: A context-based approach for pedestrian detection, Int. J. Robot. Res. (IJRR) (2013) http://dx.doi.org/10.1177/0278364912470012.
[5] J. Schlosser, C.K. Chow, Z. Kira, Fusing LIDAR and images for pedestrian detection using convolutional neural networks, in: 2016 IEEE International Conference on Robotics and Automation, ICRA, 2016, http://dx.doi.org/10.1109/ICRA.2016.7487370.
[6] Y. Tian, Z.C. Zhao, Y.J. Ma, W.P. Ding, Z.H. Gao, G.H. Cheng, L.Y. He, X.R. Zhao, Survey on deep learning in multimodal medical imaging for cancer detection, Neural Comput. Appl. (2023).
[7] P. Henry, M. Krainin, E. Herbst, X. Ren, D. Fox, RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments, Experiment. Robot. (2014).
[8] T. Shan, B. Englot, C. Ratti, D. Rus, LVI-sam: Tightly-coupled lidar-visual-inertial odometry and mapping, in: IEEE International Conference on Robotics and Automation, ICRA, 2021, http://dx.doi.org/10.1109/ICRA48506.2021.9561996.
[9] H. Gao, B. Cheng, J. Wang, K. Li, J. Zhao, D. Li, Object classification using CNN-based fusion of vision and LIDAR in autonomous vehicle environment, IEEE Trans. Ind. Inform. (2018) http://dx.doi.org/10.1109/TII.2018.2822828.
[10] X. Zhao, P. Sun, Z. Xu, H. Min, H. Yu, Fusion of 3D LIDAR and camera data for object detection in autonomous vehicle applications, IEEE Sens. J. (2020) http://dx.doi.org/10.1109/JSEN.2020.2966034.
[11] S. Hoque, M. Arafat, S. Xu, A. Maiti, Y. Wei, A comprehensive review on 3D object detection and 6D pose estimation with deep learning, IEEE Access (2021) http://dx.doi.org/10.1109/ACCESS.2021.3114399.
[12] Y. Tian, G.T. Jiang, J.W. Wang, H. Chen, L. Pan, Z.C. Xu, J.Y. Li, R.L. Wang, A revised approach to orthodontic treatment monitoring from oralscan video, IEEE J. Biomed. Health Inf. (2023).

[13] Y. Tian, H.S. Fu, H. Wang, Y.Q. Liu, Z.C. Xu, H. Chen, J.Y. Li, T.L. Wang, RGB oralscan video-based orthodontic treatment monitoring, Sci. China Inf. Sci. (2024).
[14] Q. Zhang, R. Pless, Extrinsic calibration of a camera and laser range finder (improves camera calibration, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2004, http://dx.doi.org/10.1109/IROS.2004.1389752.
[15] G. Pandey, J. McBride, S. Savarese, R. Eustice, Extrinsic calibration of a 3D laser scanner and an omnidirectional camera, IFAC Proc. Vol. (2010).
[16] W. Wang, K. Sakurada, N. Kawaguchi, Reflectance intensity assisted automatic and accurate extrinsic calibration of 3D LiDAR and panoramic camera using a printed chessboard, Comput. Vision Pattern Recognit. (2017) http://dx.doi.org/10.3390/rs9080851.
[17] J. Huang, J. Grizzle, Improvements to target-based 3D LiDAR to camera calibration, Robotics (2020) http://dx.doi.org/10.1109/ACCESS.2020.3010734.
[18] J. Cui, J. Niu, Z. Ouyang, Y. He, D. Liu, ACSC: Automatic calibration for non-repetitive scanning solid-state LiDAR and camera systems, Comput. Vision Pattern Recognit. (2020) http://dx.doi.org/10.48550/arXiv.2011.08516.
[19] J. Beltran, C. Guindel, F. Garcia, Automatic extrinsic calibration method for LiDAR and camera sensor setups, joint camera intrinsic and LiDAR-camera extrinsic calibration, IEEE Trans. Intell. Transp. Syst. (2020) http://dx.doi.org/10.1109/TITS.2022.3155228.
[20] Y. Park, S. Yun, C. Won, K. Cho, K. Um, S. Sim, Calibration between color camera and 3d lidar instruments with a polygonal planar board, Sensors (2014).
[21] F.M. Mirzaei, D.G. Kottas, S.I. Roumeliotis, 3D lidar-camera intrinsic and extrinsic calibration: Identifiability and analytical least-squares-based initialization, Int. J. Robot. Res. (2012).
[22] A. Geiger, F. Moosmann, O. Car, B. Schuster, Automatic camera and range sensor calibration using a single shot, in: IEEE International Conference on Robotics and Automation, 2012, http://dx.doi.org/10.1109/ICRA.2012.6224570.
[23] L. Zhou, Z. Li, M. Kaess, Automatic extrinsic calibration of a camera and a 3D LiDAR using line and plane, in: IEEE/RJS International Conference on Intelligent RObots and Systems, 2018, http://dx.doi.org/10.1109/IROS.2018.8593660.
[24] A. Singandhupe, H. La, Q. Ha, Single frame lidar-camera calibration using registration of 3D planes, in: IEEE International Conference on Robotic Computing, IRC, 2022, http://dx.doi.org/10.1109/IRC55401.2022.00076.
[25] Z. Zhang, A flexible new technique for camera calibration, IEEE Trans. Pattern Anal. Mach. Intell. (2000) http://dx.doi.org/10.1109/34.888718.
[26] G. Falcão, N. Hurtós, J. Massich, Plane-based calibration of a projector-camera system, Instrument. Eng. (2008).
[27] J. Weng, P. Cohen, M. Herniou, Camera calibration with distortion models and accuracy evaluation, IEEE Trans. Pattern Anal. Mach. Intell. (1992) http://dx.doi.org/10.1109/34.159901.
[28] Z. Zhang, Camera calibration with one-dimensional objects, 2004, http://dx.doi.org/10.1109/TPAMI.2004.21.
[29] C. Reich, R. Ritter, J. Thesing, White light heterodyne principle for 3D-measurement, in: Sensors, Sensor Systems, & Sensor Data Processing, International Society for Optics and Photonics, 1997, http://dx.doi.org/10.1117/12.287750.
[30] N. Schneider, F. Piewak, C. Stiller, U. Franke, Regnet: Multimodal sensor registration using deep neural networks, in: 2017, IEEE Intelligent Vehicles Symposium, IV, 2017, http://dx.doi.org/10.1109/IVS.2017.7995968.
[31] J. Shi, Z. Zhu, J. Zhang, R. Liu, Z. Wang, S. Chen, H. Liu, Calibrcnn: Calibrating camera and lidar by recurrent convolutional neural network and geometric constraints, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2020, http://dx.doi.org/10.1109/IROS45743.2020.9341147.
[32] G. Zhao, J. Hu, S. You, C.J. Kuo, Calibdnn: multimodal sensor calibration for perception using deep neural networks, signal processing, sensor/information fusion, and target recognition XXX, 2021, http://dx.doi.org/10.48550/arXiv.2103.14793.
[33] X. Lv, B. Wang, Z. Dou, D. Ye, S. Wang, Lccnet: Lidar and camera self-calibration using cost volume network, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, http://dx.doi.org/10.1109/CVPRW53098.2021.00324.
[34] L. Zhou, G. Huang, Y. Mao, S. Wang, M. Kaess, EDPLVO: Efficient direct point-line visual odometry, 2022, http://dx.doi.org/10.1109/ICRA46639.2022.9812133.
[35] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, D. Rus, LIO-SAM: Tightly-coupled lidar inertial odometry via smoothing and mapping, Robotics (2020) http://dx.doi.org/10.1109/IROS45743.2020.9341176.
[36] W. Xu, F. Zhang, FAST-LIO:.A. Fast, Robust LiDAR-inertial odometry package by tightly-coupled iterated Kalman filter, Robotics (2021) http://dx.doi.org/10.1109/LRA.2021.3064227.
[37] C. Yuan, X. Liu, X. Hong, F. Zhang, Pixel-level extrinsic self calibration of high resolution LiDAR and camera in targetless environments, Robotics (2021).