

Unlocking Wide-FoV Perception: A Robust Targetless Sensor Calibration Framework for Fisheye–LiDAR Fusion

Weijie Zhu , Shuo Shan , Kanjian Zhang , and Haikun Wei 

Abstract—Driven by the increasing accessibility of high-performance sensors, 3-D scene reconstruction is transitioning from vision-centric neural representations toward multimodal LiDAR–camera fusion to ensure absolute metric consistency. However, a critical structural bottleneck remains: existing LiDAR–camera calibration (LCC) methods rarely prioritize hardware scalability and field-of-view compatibility. Although fisheye lenses provide optimal data utilization for 360° scanning LiDARs, their severe optical distortion disrupts traditional cross-modal alignment, causing conventional geometric and statistical paradigms to fail. To bridge this gap, we propose a robust, targetless calibration framework specifically tailored for wide-FoV Fisheye–LiDAR systems. Our approach introduces a cross-modal alignment strategy that leverages LiDAR low-reflectivity edge priors and adaptive distortion modeling, achieving precise 2D-to-3D registration without physical targets. The accuracy of the method is demonstrated through extensive experiments and further validated via dynamic scene reconstruction using the FAST-LIVO2 system. The results indicate that this framework provides a reliable, on-the-fly LCC method to support omnidirectional perception hardware in practical applications.

Index Terms—3-D reconstruction, calibration, fisheye–LiDAR fusion, targetless, wide field of view (FoV).

I. INTRODUCTION

THE perception capabilities of autonomous agents set the performance ceiling for downstream tasks, such as object detection [1], place recognition [2], semantic mapping [3], and 3-D reconstruction [4]. With increasing task complexity and stricter safety requirements, LiDAR–IMU–Visual (LIV) fusion has become a predominant solution, leveraging accurate depth measurements from LiDAR, high-frequency pose estimation from IMUs, and rich semantic information from cameras [5].

Received 16 March 2026; revised 1 April 2026; accepted 9 April 2026. Date of publication 22 June 2026; date of current version 5 August 2026. This work was supported in part by the Natural Science Foundation of Jiangsu Province, China under Grant BK20241287, in part by the Fundamental Research Funds for the Central Universities under Grant 2242024k30037, and in part by the Postdoctoral Fellowship Program of Chinese Postdoctoral Science Foundation under Grant GZC20240257. Paper no. TII-26-2316. (Corresponding author: Haikun Wei.)

The authors are with the Key Laboratory of Measurement and Control of Complex Systems of Engineering, School of Automation, Southeast University, Nanjing 210096, China (e-mail: zwj@seu.edu.cn; shuoshan@seu.edu.cn; kjzhang@seu.edu.cn; hkwei@seu.edu.cn).

Digital Object Identifier 10.1109/TII.2026.3683378

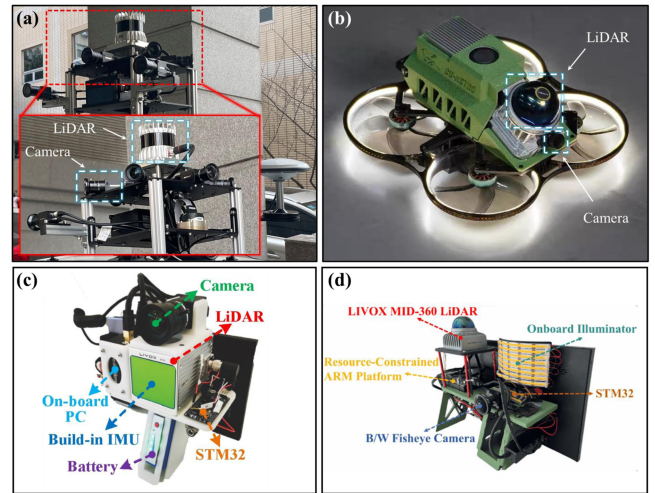


Fig. 1. Impact of hardware FoV compatibility on sensor fusion. (a) and (b) Suboptimal configurations with severely restricted FoV overlap due to conventional camera choices. (c) and (d) Optimal wide-FoV Fisheye–LiDAR setups [9] that maximize structural correspondence for 360° perception.

Deployed on platforms ranging from uncrewed vehicles [6], [7] to handheld scanning devices [8], the effectiveness of such systems fundamentally hinges on accurate LCC, which ensures spatial alignment and reliable cross-modal data association.

Despite significant progress in LCC algorithms, a critical structural bottleneck remains largely overlooked: the hardware configuration and sensor field-of-view (FoV) overlap. Algorithm designs rarely prioritize hardware scalability, leading to suboptimal sensor pairings in practice [see Fig. 1(a) and (b)]. In recent years, 360° scanning LiDARs (e.g., Livox Mid-360, Ouster) have become the industry standard. To match this omnidirectional perception, relying on standard pinhole cameras requires complex multicamera arrays [e.g., Fig. 1(a)] or results in severely limited FoV overlap [see Fig. 1(b)], restricting systems to LiDAR–inertial odometry (LIO) [10] rather than full LIV fusion [11]. In contrast, wide-FoV fisheye lenses inherently provide the most optimal FoV fusion capability and data utilization for 360° LiDARs [see Fig. 1(c) and (d)]. Unfortunately, the severe optical distortion of fisheye lenses introduces significant geometric complexities, and existing LCC methods comprehensively struggle to support this optimal hardware setup.

Historically, LCC methods have evolved through target-based, learning-based, and targetless paradigms, all of which exhibit fundamental limitations when applied to wide-FoV Fisheye–LiDAR fusion. Early target-based methods [12], [13], [14] achieve high precision using predefined patterns (e.g., checkerboards) to extract geometric constraints. However, in Fisheye–LiDAR systems, severe wide-angle distortion can distort structural features and invalidate key geometric assumptions, including edge straightness and parallelism. As a result, target-based methods may lose accuracy and become unsuitable for rapid on-site recalibration after sensor disturbance. Similarly, while learning-based approaches [15], [16] estimate extrinsics efficiently without targets, they are typically heavily overfitted to standard datasets (e.g., KITTI [17]) with narrow-FoV pinhole setups, resulting in extremely poor generalization when deployed on customized wide-FoV hardware.

To eliminate the dependence on artificial targets and specific environments, targetless LCC methods leveraging natural scene structures have become the mainstream focus [18]. However, robust cross-modal feature extraction remains a significant challenge in unstructured environments, particularly when severe wide-angle distortion is present. Existing targetless paradigms generally fail in Fisheye–LiDAR scenarios for three main reasons,

- 1) *Mutual information (MI)-based methods* [19] rely on strong correlations between image intensity and LiDAR reflectance. The pronounced grayscale inconsistency and severe exposure variations inherent to wide-FoV fisheye images break this statistical assumption, leading to calibration failure.
- 2) *Semantic-based methods* [20], [21] associate features using segmentation models, but the scale mismatch and representation differences between dense 2-D panoramic images and sparse 3-D point clouds result in weakly coupled associations.
- 3) *Geometric-based methods* [22] focus on edge alignment, yet traditional voxel-based edge extraction methods produce excessive noise at leaf node intersections, which is further exacerbated by fisheye distortion, ultimately leading to registration divergence.

To bridge this critical gap, we propose a high-precision, robust, and targetless extrinsic LCC framework tailored specifically for wide-FoV Fisheye–LiDAR fusion systems, while inherently maintaining compatibility with standard narrow-FoV pinhole setups. Our framework introduces a novel cross-modal alignment strategy that integrates prior information from the low-reflectivity edge distribution of LiDAR point clouds with tailored, adaptive distortion models for camera configurations. This approach not only eliminates the dependence on physical calibration targets but fundamentally resolves the feature alignment challenges caused by severe fisheye distortion.

Furthermore, while traditional LCC algorithms predominantly evaluate performance on single-frame static scenes, the ultimate objective of sensor fusion is to support the continuous dynamic reconstruction of the physical world. Therefore, we

directly apply our self-calibrated parameters as preset inputs for the state-of-the-art FAST-LIVO2 [23] system. Through comprehensive qualitative and quantitative analyses across various dynamic indoor and outdoor scanning reconstructions, we rigorously verify our calibration accuracy. To the best of the authors' knowledge, this work represents one of the initial attempts to validate LCC precision directly through dynamic scene reconstruction performance. The main contributions of this paper are summarized as follows.

- 1) A targetless extrinsic LCC framework strictly designed to unlock the optimal data utilization of wide-FoV Fisheye–LiDAR fusion, while maintaining full backward compatibility with standard pinhole cameras. By eliminating calibration boards, it provides a highly flexible solution for practical, on-the-fly deployments.
- 2) A novel cross-modal feature alignment strategy that effectively integrates LiDAR low-reflectivity edge priors with adaptive optical distortion models. This mathematically resolves severe fisheye geometric mismatches, enabling highly accurate 2D-to-3D edge feature association.
- 3) Comprehensive validation of the proposed framework through both static evaluations and dynamic scene reconstructions using the FAST-LIVO2 system [23]. Extensive real-world deployments across multiple indoor and outdoor scenarios demonstrate its robustness and superior performance in practical fusion applications.

II. METHODOLOGY

A. System Overview

To address the limitations of insufficient FoV overlap in conventional calibration, we propose a targetless extrinsic LCC framework tailored for wide-FoV perception. As illustrated in Fig. 2, the proposed pipeline consists of three primary stages: feature extraction, geometric association, and robust pose optimization.

First, the system processes parallel data streams from the respective sensors. Camera images are fed into a transformer-based MuGE network [24] to extract continuous 2-D edge features, which are subsequently undistorted using either a KB model [25] for fisheye lenses or a standard radial–tangential model for pinhole lenses. Concurrently, the raw LiDAR point cloud is filtered by a custom low-intensity perception (LIP) module to extract reliable 3-D structural edges.

Second, the extracted 3-D LiDAR edges are projected onto the 2-D image plane using the initial extrinsic parameters. To establish accurate cross-modal correspondences and reject outliers, the neighbor search process is strictly governed by two geometric validation criteria: a proximity constraint ($D_i^{LC} \leq \tau_d$) and an orthonormal constraint ($|n_i^C \cdot n_j^L| \geq \tau_\theta$).

Finally, to compute the optimal extrinsic parameters, we bypass traditional PnP methods in favor of a robust EPnP-based solver. The matched features undergo control point initialization to obtain an algebraic solution, which is then nonlinearly refined by minimizing a point-to-line residual coupled with a Huber loss function $\rho(\cdot)$. This ensures highly accurate and stable calibration

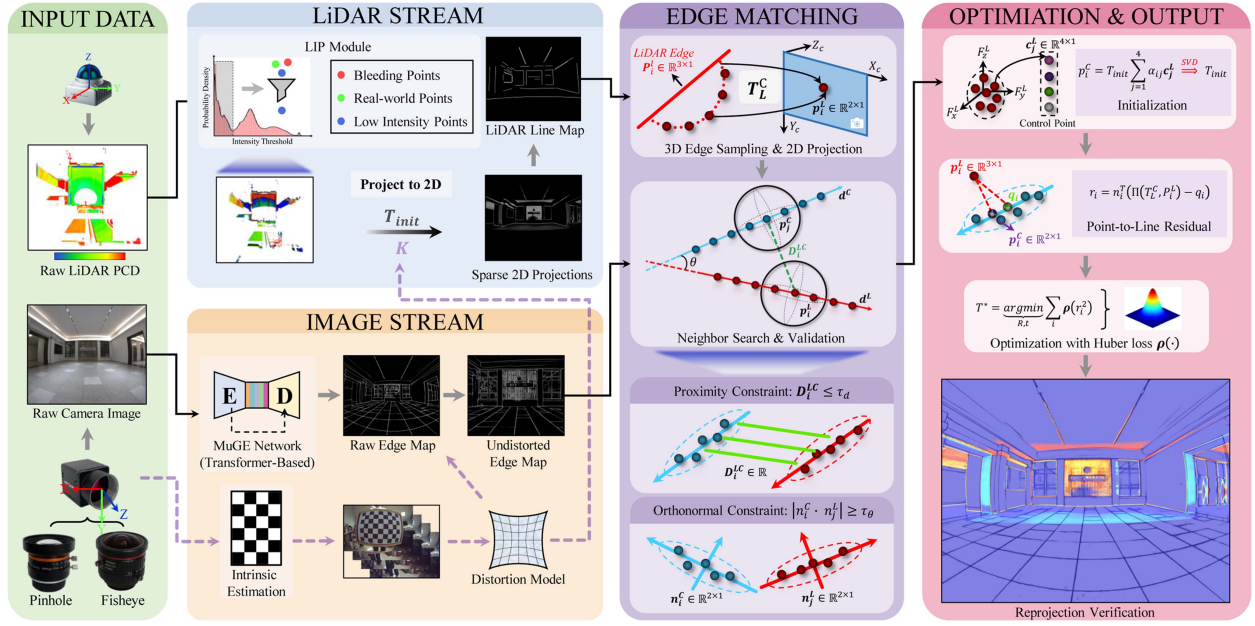


Fig. 2. System overview of the proposed targetless Fisheye–LiDAR calibration framework.

even in the presence of severe optical distortion and measurement noise.

B. LiDAR Point Cloud Feature Extraction

1) *LiDAR Intensity Perception Principle*: LiDAR reflectivity (intensity) is an intrinsic measurement attribute that exhibits high sensitivity to spatial and material variations, making it highly effective for structural edge extraction. As illustrated in the architectural scene in Fig. 3(a), intensity values fluctuate significantly at physical boundaries. For instance, distinct material transitions (e.g., concrete, glass, and metal at local feature A) yield clear intensity gradients. Similarly, sharp depth discontinuities (e.g., the structural edges of the column at local feature B) cause a severe drop in the returned intensity.

These fluctuations stem from the physical footprint of the emitted near-infrared laser spot (typically 905 nm). When the laser beam strikes a target boundary, partial occlusion and beam divergence occur, leading to reduced return energy and spatial artifacts [22]. Based on these physical optical characteristics, we formulate the low-intensity perception (LIP) principle, conceptually detailed in Fig. 3(b). This principle systematically categorizes the raw LiDAR point cloud into three distinct types.

- 1) *Real-world points*: Valid LiDAR points generated where the laser spot fully contacts the target and returns the laser signal without significant intensity loss.
- 2) *Low-intensity perceptions*: Valid but attenuated measurements located exactly at material boundaries, color transitions, or foreground structural edges. These points return lower intensity due to partial beam scattering at the edge but remain spatially accurate.
- 3) *Bleeding points*: Erroneous artifact points generated at severe depth discontinuities. When the physical laser spot

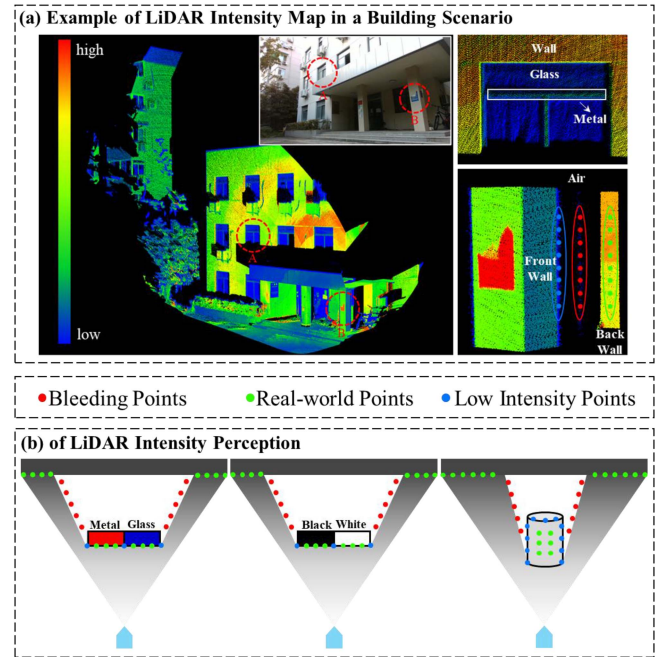


Fig. 3. (a) LiDAR intensity variations in an architectural scene. (b) Proposed LIP principle, illustrating how physical boundaries (color, material, and depth) dictate the categorization of LiDAR returns.

splits between a foreground edge and a distant background, the sensor calculates an incorrect intermediate depth, creating false “ghost” points floating in the empty space between the two surfaces.

2) *Mathematical Modeling of LiDAR Intensity*: To algorithmically implement the LIP principle and adaptively separate low intensity points from real-world points, we model the probability density of the raw LiDAR intensity distribution.

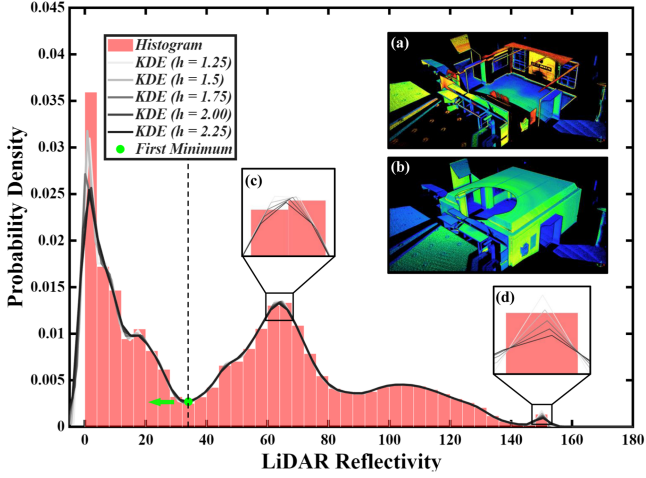


Fig. 4. Determination of the intensity threshold L via KDE under varying bandwidths (h), effectively extracting (a) the geometric skeleton from (b) the raw point cloud.

Given the inherently discrete and noisy nature of raw intensity histograms, Gaussian kernel density estimation (KDE) is employed to reconstruct a smooth, continuous probability density function (PDF)

$$\hat{f}(x) = \frac{1}{Nh\sqrt{2\pi}} \sum_{i=1}^N \exp\left(-\frac{(x-x_i)^2}{2h^2}\right) \quad (1)$$

where $\hat{f}(x)$ denotes the estimated probability density at intensity value x , $\{x_i\}_{i=1}^N$ represents the set of raw intensity observations from the point cloud, N is the total number of points, and h is the kernel bandwidth that controls the smoothing scale. As shown in Fig. 4, the 30-s static integration of a Livox Mid-360 in an indoor scene reveals a multimodal intensity distribution. Low-intensity perceptions, mainly resulting from beam scattering at structural edges and depth discontinuities, form a distinct cluster at the lower end of the spectrum. Consequently, we mathematically define the adaptive extraction threshold, L , as the *first local minimum* of the estimated PDF $\hat{f}(x)$. Points satisfying $x < L$ are retained as valid structural edge candidates. In the illustrated case, the KDE accurately localizes this physical boundary at $L = 33.5$. As shown in Fig. 4(a), applying this threshold successfully isolates the geometric skeleton of the environment from the original state, as shown in Fig. 4(b). Note that this intensity-based skeletal extraction fundamentally assumes macroscopic structural continuity, making it highly effective for typical urban and indoor scenes, though naturally less suited for highly fragmented environments (e.g., dense foliage) where continuous geometric features are inherently sparse.

Importantly, the stability of the first local minimum is strongly governed by the kernel bandwidth h . An overly small bandwidth yields an undersmoothed density estimate with spurious local minima, which can mislead threshold selection. By contrast, an excessively large bandwidth oversmooths the distribution and may eliminate the critical valley separating the low- and high-intensity components. To establish a universally robust hyperparameter, we conducted extensive evaluations across diverse scenarios. As detailed in the zoomed-in regions of Fig. 4(c)

and (d), varying h from 1.25 to 2.25 (in steps of 0.25) yields highly stable probability density curves where the location of the first minimum remains robust. Therefore, we empirically recommend $h \in [1.25, 2.25]$ as the optimal operational range. This validated range ensures the reliable extraction of the intensity threshold without the need for repetitive per-scene parameter tuning, guaranteeing the generalizability of the LIP module.

3) Cross-Modal Geometric Projection: With the 2-D image edge features extracted via the MuGE network and the 3-D structural skeleton isolated via the LIP module, the framework fundamentally relies on establishing geometric correspondences. To bridge the dimensional gap between the 3-D spatial coordinates and the 2-D pixel domain, the extracted LiDAR point cloud must be rigorously projected onto the camera image plane.

Let the filtered subset of 3-D LiDAR edge points be denoted as $\mathcal{P}^L = \{P_i^L \in \mathbb{R}^3\}_{i=1}^M$, where M is the number of valid structural points. First, these points are transformed from the LiDAR coordinate system to the camera coordinate system using the initial extrinsic parameters $\mathbf{T}_{\text{init}} = [\mathbf{R}_{\text{init}} | \mathbf{t}_{\text{init}}] \in SE(3)$

$$P_i^C = \mathbf{R}_{\text{init}} P_i^L + \mathbf{t}_{\text{init}} = [X_i^C, Y_i^C, Z_i^C]^T. \quad (2)$$

Next, the transformed 3-D points are mapped to the normalized image plane and converted into pixel coordinates using the precalibrated camera intrinsic matrix $\mathbf{K} \in \mathbb{R}^{3 \times 3}$

$$\tilde{p}_i^L = \frac{1}{Z_i^C} \mathbf{K} P_i^C \in \mathbb{R}^2 \quad (3)$$

where Z_i^C represents the depth of the point in the camera coordinate system.

At this stage, the normalized 2-D coordinates $\tilde{p}_i^L$ assume an ideal pinhole projection. To account for the severe optical distortion inherent in wide-FoV perception, a nonlinear distortion model $\mathcal{D}(\cdot)$ must be applied to obtain the final distorted pixel coordinates

$$p_i^L = \mathcal{D}(\tilde{p}_i^L) \quad (4)$$

where $\mathcal{D}(\cdot)$ represents the Kannala–Brandt (KB) model [25] for fisheye lenses or the standard radial–tangential model for pinhole lenses. This rigorous projection process establishes a unified 2-D spatial domain, serving as the foundation for the subsequent geometric feature matching and pose optimization.

4) LiDAR Edge Feature Extraction: Crucially, because the proposed LIP module has already isolated the structural skeleton from the dense raw point cloud, the projected 2-D coordinates inherently form a sparse, high-contrast, and largely noise-free edge map. This prefiltering step significantly relaxes the computational burden on the subsequent 2-D feature extraction. To parameterize these discrete projected points into continuous geometric features, we employ the highly efficient EDLines algorithm [26]. Operating on our prerefined edge map, the algorithm bypasses the computationally expensive dense gradient computations typically required for raw images. It directly links adjacent projected pixels into contiguous segments and fits geometric lines via least-squares optimization.

To guarantee robustness against any residual projection noise and prevent over detection, the extraction relies on the a contrario statistical framework inherent to EDLines. A fitted line segment of length n is only validated if its number of false alarms (NFA)

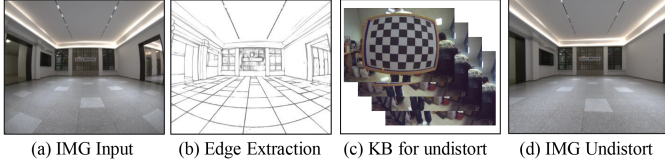


Fig. 5. Key stages of image feature processing. (a) Original input image. (b) Edge extraction result using MuGE [24]. (c) Checkerboard image captured for KB distortion parameter estimation. (d) Image after fisheye distortion removal.

satisfies $\text{NFA}(n, k) \leq 1$. This parameter-free validation ensures that the expected number of false detections under a random noise model is strictly minimized. Consequently, this highly efficient pipeline swiftly converts the discrete LiDAR projections into robust, parameterized 2-D line segments, establishing an optimal geometric foundation for the subsequent cross-modal matching with the MuGE-extracted camera edges.

C. Fisheye Image Processing

1) Fisheye Edge Extraction: In LiDAR–camera fusion, precise image edge extraction is a prerequisite for robust cross-modal feature alignment. Traditional gradient-based detectors (e.g., Canny) often struggle with hyperparameter sensitivity and typically yield fragmented boundaries, which is particularly problematic under the severe optical distortion of fisheye lenses. To overcome these limitations and ensure high-quality feature continuity, we employ a state-of-the-art transformer-based edge detection network, MuGE [24].

Rather than relying on local intensity gradients, this learning-based approach leverages global contextual information to predict edge granularity. By optimizing edge representations across both spatial and frequency domains, it effectively preserves fine structural details while suppressing noise. As illustrated in Fig. 5(b), the transformer-based extraction yields crisp, continuous, and evenly distributed edge features. This high-fidelity contour mapping provides highly reliable 2-D pixel coordinates, serving as a robust foundation for the subsequent geometric feature association with the projected LiDAR point clouds.

2) Fisheye Edge Undistortion: After extracting edge features from fisheye images, the next step is to address the distortion introduced by the fisheye lens. Due to the wide-FoV and barrel distortion inherent to fisheye lenses, the extracted edge pixels must be mapped to an undistorted coordinate system to ensure accurate matching with LiDAR data. This is achieved through the fisheye KB distortion model [25], which is used to correct the distortions and restore the image geometry to a more natural, perspective-like form. The undistortion process begins by first normalizing the distorted pixel coordinates (u_d, v_d) to a distortion plane. The pixel coordinates are transformed using the camera intrinsics f_x and f_y and the distortion parameters α , c_x , and c_y

$$\begin{bmatrix} x'_d \\ y'_d \end{bmatrix} = \begin{bmatrix} (u_d - c_x) / f_x - \alpha (v_d - c_y) / f_y \\ (v_d - c_y) / f_y \end{bmatrix} \quad (5)$$

where $r_d = \sqrt{(x'_d)^2 + (y'_d)^2}$ and $\theta_d = r_d$. The distortion angle is given by

$$\theta_d = \theta (1 + k_1\theta^2 + k_2\theta^4 + k_3\theta^6 + k_4\theta^8) \quad (6)$$

where $\theta = \arctan(r)$, with $r = \sqrt{a^2 + b^2}$. Since this relation cannot be inverted analytically, we solve for θ iteratively using Newton's method, starting with $\theta^{(0)} = \theta_d$. Once the undistorted angle θ is obtained, the normalized coordinates are restored

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{\tan \theta}{\theta_d} \begin{bmatrix} x'_d \\ y'_d \end{bmatrix}, \quad (\theta_d > 0). \quad (7)$$

Finally, the undistorted point is reprojected back onto the pixel plane (u, v)

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x(a + \alpha b) + c_x \\ f_y b + c_y \end{bmatrix}. \quad (8)$$

Before distortion correction, a checkerboard image [as shown in Fig. 5(c)] is captured to obtain the KB parameters $K = [k_1, k_2, k_3, k_4]$ required for undistortion. This undistortion process ensures that the edges extracted from fisheye images align well with the perspective geometry, providing more accurate and stable feature matching for subsequent LiDAR–camera fusion tasks. By correcting the fisheye distortion, the system is able to improve the precision of the LiDAR–camera alignment, resulting in more reliable sensor fusion performance in real-world applications.

D. Edge Matching and Optimization

With the 3-D LiDAR structural edges projected onto the 2-D image plane, the final stage of the framework establishes precise cross-modal correspondences to optimize the extrinsic parameters $\mathbf{T}_L^C$.

To geometrically associate these modalities and reject ambiguous matches, we implement a strict geometry-constrained neighbor search. Let $\mathbf{q}_i \in \mathbb{R}^2$ denote a camera edge pixel with a local normal vector $\mathbf{n}_i^C$, and let $\mathbf{p}_j^L \in \mathbb{R}^2$ denote a projected LiDAR edge point with a normal vector $\mathbf{n}_j^L$ computed from its local neighborhood. For each LiDAR point, we search for the nearest camera edge pixel. A correspondence pair $\{\mathbf{p}_j^L, \mathbf{q}_i\}$ is only considered valid if it satisfies two rigorous geometric criteria.

- 1) *Proximity constraint:* The Euclidean pixel distance between the features must fall within a threshold: $D_{ij}^{LC} = \|\mathbf{q}_i - \mathbf{p}_j^L\|_2 \leq \tau_d$.
- 2) *Orthonormal constraint:* The structural orientations of the edges must align, enforced by the dot product of their normal vectors: $|\mathbf{n}_i^C \cdot \mathbf{n}_j^L| \geq \tau_\theta$.

Once valid 2D–3D geometric associations are established, we solve for the camera pose. Conventional Perspective-n-Point (PnP) methods minimizing point-to-point reprojection errors are suboptimal for edge features (where points can slide along the contour) and are highly sensitive to mismatches. Therefore, we utilize an efficient PnP (EPnP) solver coupled with a robust point-to-line metric.

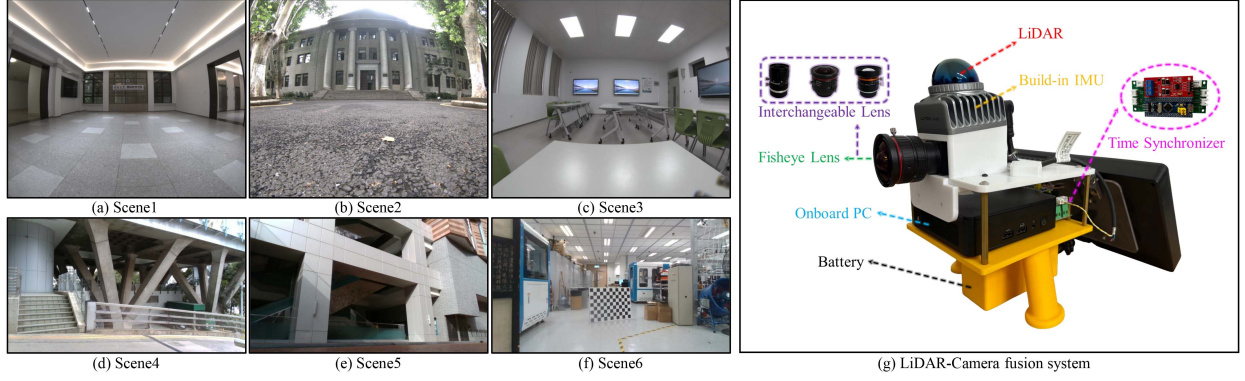


Fig. 6. Experimental scenarios and hardware setups. (a)–(c) Scenes captured with the handheld LiDAR-camera fusion system shown in (g). (d)–(f) Calibration dataset provided by Yuan et al. [22], collected using the Livox Avia sensor with locally dense point clouds and Intel RealSense-D435i.

First, the EPnP algorithm provides a rapid algebraic initialization by reparameterizing the 3-D coordinates of the matched LiDAR points ${}^C P_i$ as a weighted sum of four virtual control points $\mathbf{c}_k$

$${}^C P_i = \sum_{k=1}^4 \alpha_{ik} \mathbf{c}_k^c \xrightarrow{\text{SVD}} \mathbf{T}_{\text{init}}. \quad (9)$$

Solving this linear system via singular value decomposition (SVD) yields a stable initial extrinsic estimate $\mathbf{T}_{\text{init}}$.

To nonlinearly refine this initialization, we define a projection operator $\Pi(\mathbf{T}, {}^L P_i) = \pi(\mathbf{R} \cdot {}^L P_i + \mathbf{t})$. The orthogonal point-to-line residual r_i between the projected LiDAR point and the corresponding camera edge line is formulated as

$$r_i = (\mathbf{n}_i^C)^\top (\Pi(\mathbf{T}, {}^L P_i) - \mathbf{q}_i). \quad (10)$$

Finally, to mitigate the impact of inevitable false correspondences (outliers) that pass the initial geometric validation, we optimize this residual using a robust Huber loss function $\rho(\cdot)$. The optimal extrinsic transformation $\mathbf{T}^*$ is obtained by minimizing the sum of the robustified orthogonal distances

$$\mathbf{T}^* = \arg \min_{\mathbf{T}} \sum_i \rho(r_i^2). \quad (11)$$

This robust refinement effectively suppresses outliers, converging to a highly accurate and stable LiDAR-camera alignment even under severe wide-FoV optical distortion.

III. EXPERIMENT

A. Experimental Setup and System Configuration

To comprehensively evaluate the accuracy, robustness, and versatility of the proposed targetless extrinsic LCC framework, we designed a series of experiments using a custom-built multimodal system [see Fig. 6(g)] and public datasets [22]. Our primary testbed pairs a surround-view Livox Mid-360 LiDAR with a Hikvision industrial camera. To rigorously compare fusion modalities and highlight the spatial advantages of wide-FoV configurations, this camera is equipped with two interchangeable lenses: a 2.8 mm fisheye (LTCAM) and a 6.0 mm standard pinhole (MVL-HF0628 M). The lens selection strictly ensures

the image circle exceeds the camera's sensor size, effectively preventing peripheral vignetting and guaranteeing high-quality feature extraction. This dual-lens setup directly validates our core Fisheye-LiDAR fusion while confirming backward compatibility with conventional Pinhole-LiDAR systems.

Beyond the custom testbed, we verify the framework's generalizability across diverse, off-the-shelf architectures using the open-source Livox Calib dataset [22], which features a Livox Avia LiDAR and an Intel RealSense D435 camera. Furthermore, to demonstrate practical deployment capabilities, we integrated our custom hardware into a handheld platform running FAST-LIVO2 [23]. This setup enables real-time mapping and precise sensor synchronization for dynamic scene reconstruction. The detailed specifications for all hardware configurations are summarized in Table I.

B. Spatial Superiority of Wide-FoV Fusion

A primary motivation for developing our targetless extrinsic LCC framework is to exploit the substantial spatial advantages of Fisheye-LiDAR fusion over conventional pinhole setups. While our method inherently supports both hardware configurations, the perceptual superiority of the wide-FoV system is readily apparent, as conceptually illustrated in Fig. 7(a). To quantitatively demonstrate this, we conducted a controlled comparison using the hardware specifications listed in Table I. By maintaining a fixed vantage point and a constant LiDAR integration time of 20 s, we captured image data for both fisheye and pinhole lenses.

Considering the focal lengths and FoV of each lens, the fisheye configuration (2.8 mm, $172.2^\circ \times 155.5^\circ$) provides a roughly fourfold horizontal and vertical coverage expansion relative to the pinhole lens (6.0 mm, $45.4^\circ \times 35.5^\circ$). Applying our LCC pipeline to both setups, the resulting color-fused point clouds, visualized in Fig. 7(b) and (c), show that the fisheye system successfully colorizes approximately twice the environmental area for identical raw LiDAR input. This increase in effective coverage translates directly to a higher number of usable edges for calibration and denser 3-D reconstruction, without incurring extra computational or hardware costs and only a simple lens substitution is required.

TABLE I
DETAILED SPECIFICATIONS OF THE CUSTOM FISHEYE-LiDAR AND
PINHOLE-LiDAR FUSION HARDWARE

Parameter	Specification
LiDAR: Livox Mid-360	
Detection Range	40 m @ 10% reflectivity 70 m @ 80% reflectivity
Scanning Frequency	10 Hz
Angular Resolution	0.75°
Field of View	360° (H) × 59° (V) (−7° ~ 52°)
Camera Body: Hikvision MV-CU050-90UC	
Resolution	3072 × 2048
Max Frame Rate	59.8 FPS
Target Sensor Size	1/1.8" (compatible with both lenses)
Interchangeable Lenses	
Pinhole Lens	MVL-HF0628M
Focal Length	6.0 mm
Supported Resolution	6 Megapixels (6MP)
Field of View	45.4° (H) × 35.5° (V)
Fisheye Lens	LTCAM
Focal Length	2.8 mm
Supported Resolution	10 Megapixels (10MP)
Field of View	172.2° (H) × 155.5° (V)

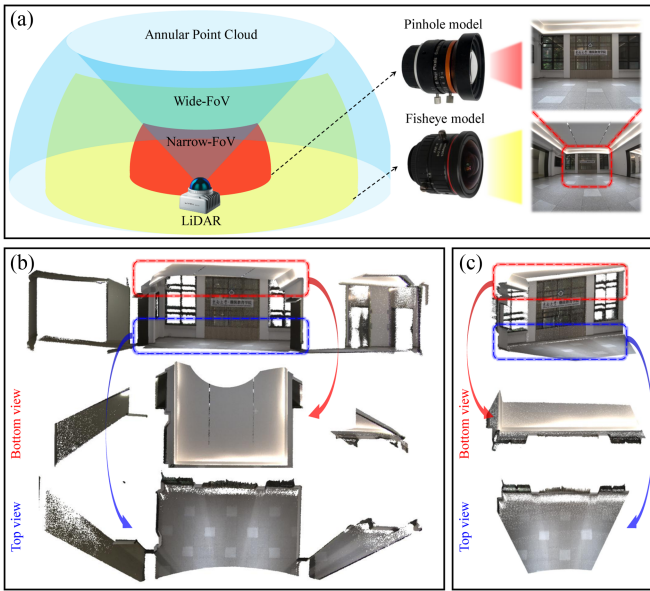


Fig. 7. Comparison of LiDAR-Camera fusion performance under wide-FoV and narrow-FoV configurations. (a) Schematic of FoV overlap in different LiDAR-camera fusion configurations. (b) Colorized point cloud with wide-FoV fusion. (c) Colorized point cloud with narrow-FoV fusion.

C. Qualitative Assessment of Calibration Precision

Having established the system-level benefits of wide-FoV fusion, we qualitatively evaluate the underlying calibration accuracy that enables it. The raw LiDAR point clouds are

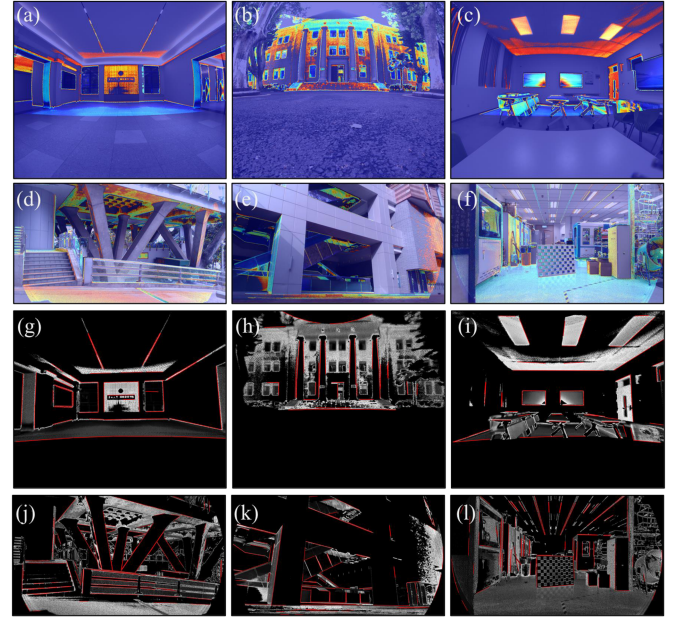


Fig. 8. LiDAR reprojection and edge map visualization for natural scenes [see Fig. 6(a)–(f)]. Subfigures (a)–(f) show the LiDAR points re-projected onto images to verify cross-modal overlap, while subfigures (g)–(l) show edges extracted from low-reflectivity LiDAR points. This visualization highlights the spatial distribution of usable points for calibration and the quality of edge extraction under diverse natural scene conditions.

reprojected onto the corresponding image planes using the extrinsic parameters estimated by our proposed LCC framework, allowing for a direct visual inspection of the global alignment. As illustrated across six diverse indoor and outdoor scenarios in Fig. 8(a)–(f), the proposed method achieves highly accurate global point-to-pixel registration. Notably, this exceptional performance remains consistent whether processing images with severe wide-angle fisheye distortion [see Fig. 6(a)–(c)] or standard narrow-FoV images with minimal distortion [see Fig. 6(d)–(f)]. This robustness and cross-hardware compatibility are fundamentally attributed to the framework’s adaptive integration of appropriate camera distortion models during the optimization process.

Furthermore, Fig. 8(g)–(l) provides a detailed visualization of the projected LiDAR edge features. Facilitated by accurate intrinsic parameters and the low-reflectivity filtering of our LIP strategy, these projected edge points are precisely distributed along the structural contours of the physical environment, particularly localizing at material transitions and depth discontinuities. This uniformly distributed, high-quality contour mapping significantly enhances the precision of cross-modal edge extraction, which in turn guarantees robust feature association and highly stable calibration. Specifically, the crisp alignment of indoor door frames and windows in Fig. 8(a)–(c), along with the perfectly overlapped 12×8 target contours in Fig. 8(f), provides compelling visual evidence confirming the high precision and overall effectiveness of the proposed Fisheye–LiDAR fusion framework.

TABLE II
QUANTITATIVE COMPARISON WITH STATE-OF-THE-ART TARGETLESS LCC METHODS ACROSS SIX SCENES

Scene	e_t (cm)					e_R (deg)				
	E2E [27]	Yuan [22]	MFCalib [28]	MI [19]	Proposed	E2E [27]	Yuan [22]	MFCalib [28]	MI [19]	Proposed
1	18.1773	40.6789	562.8807	23.6289	3.1031	5.6165	11.7615	42.0981	12.8246	4.8147
2	108.2611	38.0540	243.9871	29.6135	3.8172	15.6488	4.1293	10.4631	13.4812	4.6793
3	77.6996	29.1524	27.8424	32.7572	3.3498	7.5039	12.4386	5.6350	4.7553	4.6481
4	57.6281	17.8874	346.7258	27.4982	5.1587	8.1573	8.9114	15.3716	18.6548	3.4279
5	67.8275	23.7853	178.2489	24.7259	5.9157	8.8561	5.4112	12.8159	22.9184	5.8342
6	74.2985	21.1579	279.5928	30.5280	6.6390	17.6542	7.7561	16.3591	25.5172	5.7228
Avg.	67.3153	28.4527	273.2175	28.1253	4.6639	10.5728	8.4014	17.1235	16.3586	4.8545
Std. Dev.	26.9004	8.4536	163.0109	3.2044	1.3287	4.4474	3.0402	11.7068	6.9233	0.7983

Best results are in bold (green for individual scenes, red for statistical metrics).

D. Benchmarking Targetless Calibration Performance

As discussed in the introduction, most existing targetless LCC methods inherently assume standard pinhole projection models and lack native support for wide-FoV fisheye images. To ensure a strictly fair and rigorous quantitative comparison, the raw fisheye images are first undistorted using the KB model (as mentioned in Section II-C2) before being processed by the baseline algorithms. We benchmark our proposed framework against four state-of-the-art targetless methods, which are categorized by their core alignment mechanisms: MI-based [19], geometric edge alignment-based [22], [27], and reflectance-depth fusion-based [28] approaches. For clarity, all baseline methods were run using their respective optimal or officially recommended parameters; the same initial conditions, image resolution, and point cloud preprocessing were applied across all methods; and all methods received identically undistorted images as input, ensuring that the undistortion operation neither advantaged nor disadvantaged any baseline. These evaluations are conducted across the six diverse natural scenes depicted in Fig. 6(a)–(f).

The calibration accuracy is quantitatively assessed using two standard metrics: translation error e_t and rotation error e_R [29], defined as follows:

$$e_t = \|\hat{t} - t\| \quad (12)$$

$$e_R = \arccos \left(\frac{\text{tr}(\hat{R} \cdot \hat{R}^{-1} - 1)}{2} \right) \quad (13)$$

where $\hat{R}$ and $\hat{t}$ denote the estimated rotation matrix and translation vector, respectively, while R and t represent the corresponding ground-truth extrinsic parameters characterizing the spatial transformation between the LiDAR and camera optical centers. For the custom hardware setup employed in these experiments, the precise ground-truth parameters are physically measured from CAD and established as

$$R = \begin{bmatrix} 0 & -1.0000 & 0 \\ 0.2079 & 0 & -0.9781 \\ 0.9781 & 0 & 0.2079 \end{bmatrix}, \quad t = \begin{bmatrix} 0 \\ 0.065 \\ 0.082 \end{bmatrix}. \quad (14)$$

Based on the established ground truth and evaluation metrics, Table II summarizes the quantitative calibration errors of all

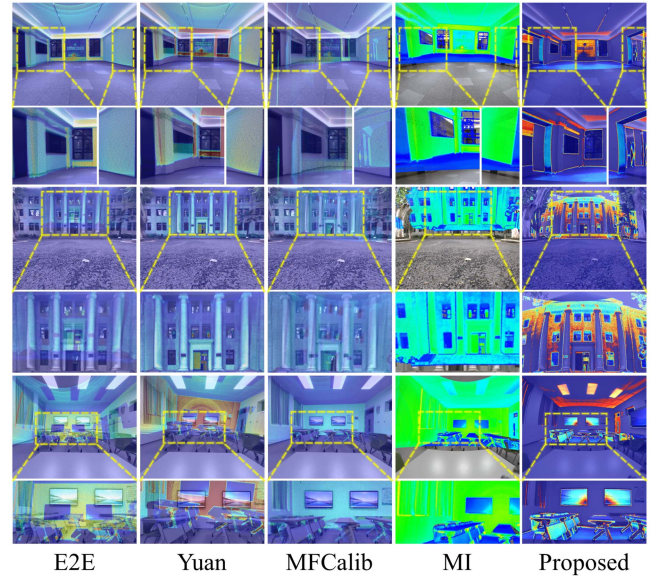


Fig. 9. LiDAR–camera calibration results obtained by the proposed method and baseline approaches: E2E [27], Yuan et al. [22], MFCalib [28], and MI [19]. Columns from left to right correspond to each method. The visualization demonstrates cross-modal alignment quality, showing that the proposed method achieves more complete and accurate feature correspondence in natural scenes.

evaluated methods across the six natural scenes. The proposed targetless extrinsic LCC framework demonstrates superior performance, achieving the lowest average rotation error of 4.85° (with a standard deviation of 0.80°) and an average translation error of 4.66 cm (with a standard deviation of 1.33 cm).

To clearly illustrate the performance difference between our approach and the baselines, Fig. 9 presents comparative reprojection visualizations for three fisheye scenarios, particularly under severe wide-angle distortion. Crucially, the quantitative metrics strongly correlate with these qualitative visual results. For instance, the E2E method exhibits significant rotation (15.65°) and translation (108.26 cm) errors in Scene 2. This numerical degradation aligns perfectly with the obvious calibration drift and edge misalignment observed in the corresponding zoomed-in reprojection map (see Fig. 9, row 4, column 1).

Ultimately, evaluated on both absolute accuracy and cross-scene stability (indicated by the minimal standard deviations),

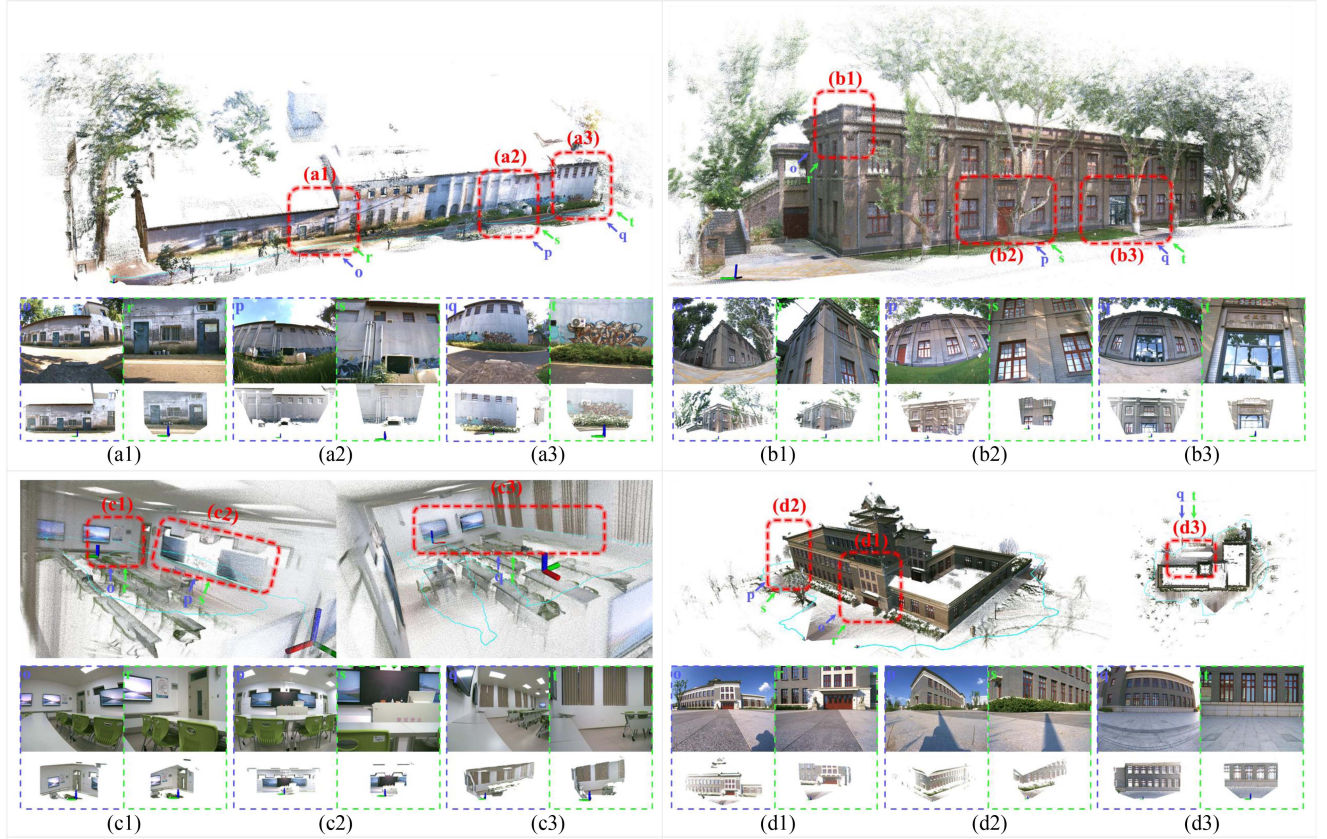


Fig. 10. Colored point cloud maps reconstructed by FAST-LIVO2 using extrinsics from the proposed wide-FoV targetless extrinsic LCC framework. (a) Long wall. (b) School history museum. (c) Classroom. (d) Ancient architecture. For each scene (a1–a3, b1–b3, c1–c3, d1–d3), results compare wide- and narrow-FoV lenses under identical integration time, where the first row shows the camera views and the second row shows the reconstructed point clouds.

the proposed method significantly outperforms existing state-of-the-art techniques. By delivering exceptional robustness across both fisheye and pinhole modalities, this framework effectively bridges the critical technical gap in targetless calibration for wide-FoV Fisheye–LiDAR fusion systems.

E. Real-World Mapping Experiments

On the handheld LiDAR–camera platform shown in Fig. 6(g), extrinsic calibration was first performed using the proposed targetless extrinsic LCC framework. To validate the feasibility of wide-FoV sensing, the FAST-LIVO2 [23] algorithm was deployed for real-world mapping, with strict hardware-level synchronization inspired by FAST-LIVO [9]. Real-time processing was executed on an ASUS NUC 13 Pro. Four representative scenarios were selected for evaluation, including a long wall, a school history museum, a classroom, and an ancient building, corresponding to Fig. 10(a)–(d). In each scenario, three viewpoints were randomly selected to conduct static scanning experiments, where mapping results obtained with narrow- and wide-FoV camera lenses were compared under the same LiDAR integration time of 20 s. All experiments were conducted after calibrating both lens settings using the proposed targetless extrinsic LCC framework.

The results provide two key insights. First, from the perspective of scene completeness, the reconstructed maps preserve local color details and structural textures, demonstrating the high calibration accuracy achieved by the proposed method. Moreover, the reconstructed point clouds across all four scenes consistently exhibit an effective depth coverage of approximately 5–30 m using the same calibrated extrinsic parameters, indicating that the calibration results generalize well across different environments. Second, by contrasting narrow- and wide-FoV lenses under identical conditions, it is evident that the wide-FoV configuration achieves superior fusion completeness of raw point clouds, thereby exploiting the full potential of LiDAR–camera fusion in extending perceptual coverage. These findings confirm both the practicality and significance of the proposed wide-FoV LiDAR–camera fusion system.

F. Ablation Study

To evaluate the effectiveness of the proposed LIP module at the feature extraction stage, we conducted ablation experiments across six scenes. As shown in Fig. 11, the BASE method, which includes all modules, achieved mean rotation and translation errors of 4.85° and 4.66 cm, respectively. Removing the LIP module increased the rotation and translation errors to 6.51°

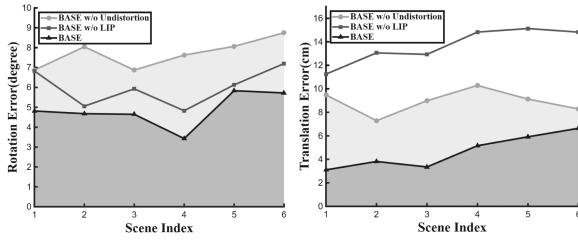


Fig. 11. Rotation error (left) and translation error (right) across six scenes for BASE, BASE without LIP, and BASE without Undistortion. Each line represents one method, illustrating the impact of removing each module on calibration accuracy.

TABLE III

COMPUTATION TIME STATISTICS FOR EACH MODULE OF THE PROPOSED METHOD, INCLUDING CEE, LEE, ROUGH CALIBRATION (FEATURE MATCHING), AND REFINED CALIBRATION (EPnP OPTIMIZATION)

Scene	Loading (s)		Core Algorithm Time (s)				Total (s)
	IMG	PCD	CEE	LEE	Rough	Refine	
scene1	0.034	4.522	0.054	0.233	1.978	4.143	10.964
scene2	0.036	4.978	0.052	0.209	2.467	4.818	12.560
scene3	0.035	5.217	0.046	0.435	2.551	5.956	14.240
scene4	0.032	11.126	0.039	0.541	6.174	11.320	29.232
scene5	0.034	11.178	0.037	0.338	3.072	6.430	21.089
scene6	0.032	11.135	0.027	0.306	2.992	6.184	20.676
Average	0.034	8.026	0.043	0.344	3.206	6.475	18.127

The total cost time for the proposed method is highlighted.

and 8.91 cm. While the rotation error increase was modest, the translation error rose by approximately 4 cm, noticeably reducing calibration accuracy. This highlights the importance of the LIP module in providing high-quality feature information to the edge extraction process.

Similarly, ablation of the Undistortion module using fisheye and slightly distorted pinhole images resulted in a significant increase in translation error, with an average of 13.67 cm. This indicates that distortion correction is even more critical than the LIP module for maintaining precise calibration. Overall, these results demonstrate that each module in the proposed framework contributes substantially to accurate LiDAR–camera calibration, with the LIP module enhancing feature extraction inputs and the Undistortion module ensuring reliable geometric alignment under distorted imaging conditions.

G. Analysis of Computational Efficiency

We evaluate the runtime of the proposed calibration pipeline on a workstation equipped with an NVIDIA GeForce RTX 4090 GPU, an Intel Core i9-13900 K CPU, and 16 GB of RAM. Per-scene computation times are summarized in Table III, which decomposes the total latency into data loading, core algorithmic modules [Camera edge extraction (CEE), LiDAR edge extraction (LEE), rough calibration via feature matching, and refined calibration via EPnP optimization), and total time cost. As evidenced by the table, point cloud loading constitutes the primary I/O bottleneck, ranging from 4.522 to 5.217 s for scenes 1–3 and reaching up to 11.178 s for scenes 4–6; this discrepancy arises from the increased point cloud density in the public dataset relative to the self-collected scenes. In contrast,

image loading incurs negligible overhead (approximately 0.03 s) across all scenarios.

Among the core processing steps, the EPnP-based refinement stage exhibits the highest computational demand, averaging 6.475 s, followed by the rough calibration stage at 3.206 s, whereas the edge extraction modules for both camera and LiDAR contribute only marginal runtime, each below 0.4 s. Total processing time varies from 10.964 to 29.232 s, with a mean of 18.127 s. Notably, scenes from the public dataset consistently incur higher latency due to their larger point cloud sizes and more intricate geometric structures, which impose heavier loads on feature matching and optimization. Overall, the pipeline demonstrates practical runtime characteristics suitable for offline calibration, with identified bottlenecks in point cloud I/O and refinement. These observations suggest straightforward avenues for acceleration, such as strategic point cloud down-sampling or parallelized EPnP computation.

IV. CONCLUSION

This article presents a targetless LCC method that supports wide-FoV perception by leveraging the FoV overlap characteristics between sensors. The proposed LiDAR–camera fusion system effectively integrates annular point clouds and fisheye images, two types of wide-FoV sensing primitives, offering insights for handling scenarios with mismatched or suboptimal sensor configurations. Extensive qualitative and quantitative comparisons with state-of-the-art targetless LCC methods show the proposed method achieves high calibration accuracy, well-suited for wide-FoV scenarios. Real-world mapping experiments validate cross-scene generalization of calibrated extrinsics, with consistent performance across diverse scenes using a single parameter set, demonstrating strong robustness and generalization under varying conditions.

REFERENCES

- [1] H. Xu, S. Huang, Y. Yang, X. Chen, and S. Hu, “Deep learning-based pedestrian detection using RGB images and sparse LiDAR point clouds,” *IEEE Trans. Ind. Informat.*, vol. 20, no. 5, pp. 7149–7161, May 2024.
- [2] J. Ma, G. Xiong, J. Xu, and X. Chen, “CVTNet: A cross-view transformer network for LiDAR-based place recognition in autonomous driving environments,” *IEEE Trans. Ind. Informat.*, vol. 20, no. 3, pp. 4039–4048, Mar. 2024.
- [3] J. S. Berrio, M. Shan, S. Worrall, and E. Nebot, “Camera–LiDAR integration: Probabilistic sensor fusion for semantic mapping,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 7637–7652, Jul. 2022.
- [4] W. Zhen, Y. Hu, J. Liu, and S. Scherer, “A joint optimization approach of LiDAR–camera fusion for accurate dense 3-D reconstructions,” *IEEE Robot. Autom. Lett.*, vol. 4, no. 4, pp. 3585–3592, Apr. 2019.
- [5] S. Li, X. Li, S. Chen, Y. Zhou, and S. Wang, “Two-step LiDAR/camera/IMU spatial and temporal calibration based on continuous-time trajectory estimation,” *IEEE Trans. Ind. Electron.*, vol. 71, no. 3, pp. 3182–3191, Mar. 2024.
- [6] W. Yang et al., “Autonomous–targetless extrinsic calibration of thermal, RGB, and LiDAR sensors,” *IEEE Trans. Instrum. Meas.*, vol. 73, 2024, Art. no. 4510011, doi: 10.1109/TIM.2024.3480237.
- [7] D. Cattaneo and A. Valada, “CMRNext: Camera to LiDAR matching in the wild for localization and extrinsic calibration,” *IEEE Trans. Robot.*, vol. 41, pp. 1995–2013, 2025.
- [8] B. Zhou, C. Zheng, Z. Wang, F. Zhu, Y. Cai, and F. Zhang, “FAST-LIVO2 on resource-constrained platforms: LiDAR–inertial–visual odometry with efficient memory and computation,” *IEEE Robot. Autom. Lett.*, vol. 10, no. 8, pp. 7931–7938, Aug. 2025.

- [9] C. Zheng, Q. Zhu, W. Xu, X. Liu, Q. Guo, and F. Zhang, "FAST-LIVO: Fast and tightly-coupled sparse-direct LiDAR-inertial-visual odometry," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2022, pp. 4003–4009.
- [10] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "FAST-LIO2: Fast direct LiDAR-inertial odometry," *IEEE Trans. Robot.*, vol. 38, no. 4, pp. 2053–2073, Apr. 2022.
- [11] L. Gong et al., "preciseSLAM: Robust, real-time, LiDAR-inertial-ultrasonic tightly-coupled SLAM with ultraprecise positioning for plant factories," *IEEE Trans. Ind. Informat.*, vol. 20, no. 6, pp. 8818–8827, Jun. 2024.
- [12] L. Zhou, Z. Li, and M. Kaess, "Automatic extrinsic calibration of a camera and a 3 D LiDAR using line and plane correspondences," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2018, pp. 5562–5569.
- [13] G. Koo, J. Kang, B. Jang, and N. Doh, "Analytic plane covariances construction for precise planarity-based extrinsic calibration of camera and LiDAR," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2020, pp. 6042–6048.
- [14] C. Zheng and F. Zhang, "FAST-Calib: LiDAR-camera extrinsic calibration in one second," 2025, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167865511001772>
- [15] G. Iyer, R. K. Ram, J. K. Murthy, and K. M. Krishna, "CalibNet: Geometrically supervised extrinsic calibration using 3 D spatial transformer networks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2018, pp. 1110–1117.
- [16] X. Lv, B. Wang, Z. Dou, D. Ye, and S. Wang, "LCCNet: LiDAR and camera self-calibration using cost volume network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, 2021, pp. 2888–2895.
- [17] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2012.
- [18] S. Li, X. Li, S. Chen, Y. Zhou, and S. Wang, "Targetless spatiotemporal calibration of multi-LiDAR multi-IMU system based on continuous-time optimization," *IEEE Trans. Ind. Informat.*, vol. 20, no. 5, pp. 7565–7575, May 2024.
- [19] G. Pandey, J. R. McBride, and R. M. Eustice, "Automatic calibration of LiDAR and camera images using normalized mutual information," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2013, pp. 613–619.
- [20] Y. Zhu, C. Li, and Y. Zhang, "Online camera-LiDAR calibration with sensor semantic information," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2020, pp. 4970–4976.
- [21] A. Kirillov et al., "Segment anything," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2023, pp. 3992–4003.
- [22] C. Yuan, X. Liu, X. Hong, and F. Zhang, "Pixel-level extrinsic self calibration of high resolution LiDAR and camera in targetless environments," *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 7517–7524, Apr. 2021.
- [23] C. Zheng et al., "FAST-LIVO2: Fast, direct LiDAR-inertial-visual odometry," *IEEE Trans. Robot.*, vol. 41, pp. 326–346, 2025.
- [24] C. Zhou, Y. Huang, M. Pu, Q. Guan, R. Deng, and H. Ling, "MuGE: Multiple granularity edge detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2024, pp. 25952–25962.
- [25] J. Kannala and S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 8, pp. 1335–1340, Aug. 2006.
- [26] C. Akinlar and C. Topal, "Edlines: A real-time line segment detector with a false detection control," *Pattern Recognit. Lett.*, vol. 32, no. 13, pp. 1633–1642, 2011. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167865511001772>
- [27] W. Zhu, S. Shan, C. Tong, K. Zhang, and H. Wei, "Targetless automatic edge-to-edge extrinsic calibration of LiDAR-camera system based on pure laser reflectivity," *IEEE Trans. Instrum. Meas.*, vol. 74, 2025, Art. no. 8507714.
- [28] T. Ye, W. Xu, C. Zheng, and Y. Cui, "MFCalib: Single-shot and automatic extrinsic calibration for LiDAR and camera in targetless environments based on multi-feature edge," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2024, pp. 864–871, doi: [10.1109/IROS58592.2024.10802395](https://doi.org/10.1109/IROS58592.2024.10802395).
- [29] S. Xing, F. Jing, and M. Tan, "Reconstruction-based hand-eye calibration using arbitrary objects," *IEEE Trans. Ind. Informat.*, vol. 19, no. 5, pp. 6545–6555, May 2023.