Direct, Targetless and Automatic Joint Calibration of LiDAR-Camera Intrinsic and Extrinsic

Yishu Shen[1*], Sheng Hong[2*], Shaojie Shen[2], and Tong Qin[1†]

***Abstract*— This paper presents a direct, targetless, and automatic LiDAR-camera joint calibration method that effectively overcomes the intrinsic precision limitations inherent in conventional vision-based techniques. We propose an iterative two-stage optimization framework that leverages 3D LiDAR measurements to simultaneously refine both intrinsic and extrinsic. In the first stage, initial extrinsic parameters derived from CAD specifications facilitate the projection of LiDAR point clouds onto the camera image plane, while a normalized information distance (NID) metric a measure from information theory that quantifies the statistical alignment between LiDAR and image intensities, is employed to optimize intrinsic. In the second stage, the refined intrinsic guides further optimization of extrinsic using the same NID-based evaluation metrics. This alternating process iteratively enhances both intrinsic and extrinsic through their mutual interdependence. Experiments across multiple datasets demonstrate that our method achieves sub-pixel intrinsic accuracy and extrinsic parameters that closely align with CAD specifications, validating the superior performance of our approach for sensor fusion applications.**

I. Introduction

Camera intrinsic calibration and LiDAR-camera extrinsic calibration are fundamental prerequisites for various robotics and computer vision applications, particularly in simultaneous localization and mapping (SLAM) systems [1, 2] and object detection [3, 4]. The accuracy of these calibrations directly impacts the performance of sensor fusion.

However, current target-based intrinsic calibration methods highly rely on camera detection of targets such as checkerboards [5]–[7]. Due to inherent camera resolution limitations, these intrinsic calibration methods are fundamentally constrained, restricting their accuracy to the sub-pixel level and making higher precision unattainable. Furthermore, camera intrinsic are susceptible to external factors such as temperature variations [8, 9], rendering previously calibrated parameters invalid, while the targetless intrinsic calibration method for recovering these parameter changes demonstrates inadequate accuracy. Current extrinsic calibration methods typically require precise projection of LiDAR 3D points onto the camera image plane [10], thus imposing stringent requirements on intrinsic parameter accuracy. Consequently, even minor intrinsic parameter inaccuracies propagate into significant extrinsic calibration errors. These errors manifest as deviations from ground-truth CAD values despite visually satisfactory 3D-2D alignments [11, 12]. Given these limitations, accurate intrinsic and extrinsic calibration methods still present two primary challenges: 1) the absence of precise targetless intrinsic calibration methods achieving sub-pixel level accuracy, and 2) extrinsic calibration accuracy being tightly influenced by imprecise intrinsic, lacking a joint method that optimizes extrinsic with refined intrinsic.

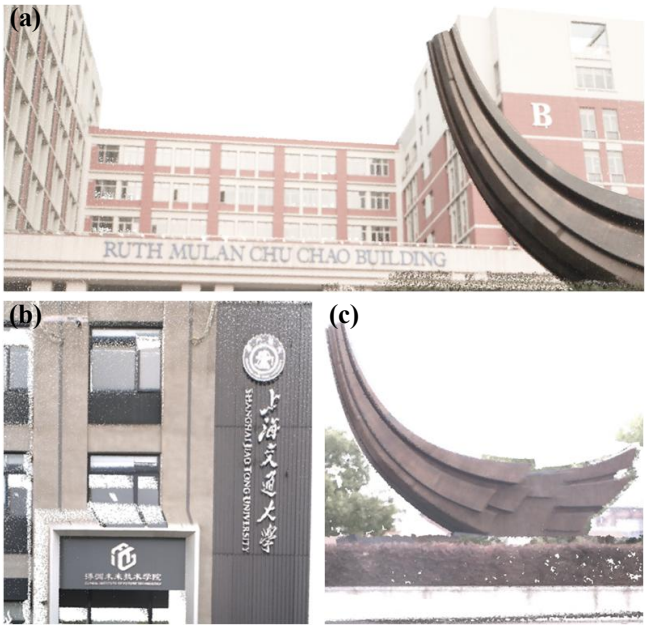


Fig. 1: Colorized visualization of LiDAR point clouds from three distinct scenarios (a)–(c), colorized through projection with intrinsic and extrinsic obtained via our proposed calibration method. The results demonstrate accurate alignment between LiDAR point clouds and camera images.

To address these challenges, we propose a direct targetless joint calibration approach. Fig. 1 presents the colorized visualization of LiDAR point clouds using the precise intrinsic and extrinsic obtained through our method. The key contributions of this work include:

- A direct calibration method utilizing NID-based evaluation criterion and LiDAR's precise 3D measurements to enhance camera intrinsic calibration beyond traditional sub-pixel limitations, while ensuring consistency with

* Equal contribution.
†Corresponding author (email: qintong@sjtu.edu.cn).
[1]Yishu Shen and Tong Qin are with the Global Institute of Future Technology, Shanghai Jiao Tong University, Shanghai, China.
[2]Sheng Hong and Shaojie Shen are with Department of Electronic Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China.
This work was supported by the Natural Science Foundation of Shanghai (Grant No. 24ZR1435600).

CAD specifications.

- An efficient two-stage optimization framework that sequentially refines intrinsic before conducting joint optimization, thereby mitigating parameter overfitting while preserving the critical interdependencies between intrinsic and extrinsic calibration.
- Extensive experimental validation demonstrating superior calibration accuracy and robustness across distinct self-integrated sensor suites and real-world scenarios.
- Open-source code of our joint calibration method to benefit the community [1].

II. Related work

A. Camera Intrinsic Calibration

Current camera calibration methods, exemplified by Zhang's seminal methodology employing the checkerboards [5], have achieved widespread adoption within the research community due to their implementation simplicity and calibration efficacy. However, these methods are fundamentally constrained to sub-pixel accuracy due to inherent resolution limitations. Recent advancements in targetless camera intrinsic calibration methods [13]–[15] have attempted to transcend the dependency on calibration targets by exploiting inherent features of the scene. However, these methods frequently exhibit performance degradation due to feature correspondence ambiguities and lack of prior depth information, yielding inferior calibration accuracy compared to target-based approaches.

For specialized camera systems such as omnidirectional cameras, more sophisticated calibration methods have been proposed [16]–[18]. However, these methods necessitate meticulous parameter optimization and exhibit vulnerability to overfitting, particularly when implementing high-order polynomial representations.

B. LiDAR-Camera Extrinsic Calibration

Extrinsic calibration methods for LiDAR-camera systems can be categorized into target-based and targetless approaches. Target-based methods [19]–[21] often rely on specifically engineered calibration targets detectable by both sensors. While these methods achieve high precision, they necessitate careful target design and manual intervention throughout the procedure [22].

Targetless calibration methods have gained substantial attention owing to their operational flexibility and practical deployment capabilities, such as checkerboards and 3D structured objects [23, 24]. These methods predominantly bifurcate into two taxonomic categories: edge-based methods [11, 25, 26] that establish correspondence through geometric feature alignment extracted from disparate modalities, and intensity-based techniques that exploit mutual information metrics between sensor data streams. Recent advancements by Koide et al. [10] have demonstrated notable calibration efficacy utilizing normalized information distance (NID) for extrinsic estimation, demonstrating robustness in 3D-2D alignment across diverse environmental contexts without necessitating explicit feature extraction procedures. However, their method fundamentally presupposes the accuracy of camera intrinsic. When intrinsic calibration precision is impaired, even calibration results exhibiting visually satisfactory alignment can generate extrinsic results that deviate significantly from manufacturer-specified CAD specifications.

C. Joint Intrinsic-Extrinsic Calibration

Given the critical impact of camera intrinsic on calibration accuracy, traditional methods have predominantly addressed camera intrinsic and LiDAR-camera extrinsic calibration as discrete optimization problems [27]. However, these decoupled approaches insufficiently address the tightly coupled interdependencies between the extrinsic and intrinsic, thereby imposing fundamental limitations on achievable calibration precision.

Building on these insights, recent studies [28, 29] have demonstrated that these parameters are inherently coupled, suggesting that their concurrent optimization may enhance calibration fidelity. Recent advancements in joint calibration methods [30, 31] have shown promising results via 3D-2D edge feature extraction and correspondence. However, these edge-based methods are highly dependent on hyperparameters, such as voxel dimensions for 3D edge detection, Canny thresholds for 2D edge identification, and nearest neighbor search criteria for correspondence establishment, which often reduce robustness in practical deployments and result in prolonged calibration processing times. Moreover, direct simultaneous optimization of all parameters in these methods may lead to overfitting and potentially yield physically implausible solutions.

To address these limitations, we propose a novel method inspired by information-theoretic metrics. Our calibration method leverages the Normalized Information Distance (NID) metric, which evaluates the statistical co-occurrence of intensity values [32, 33] rather than direct differences between LiDAR and image photometric intensities. This method satisfies the metric space axioms and demonstrates enhanced robustness [10, 34].

III. Methodology

As shown in Fig. 2, our calibration method comprises several key steps. Initially, we integrate the sensor suites consisting of LiDAR and camera. Subsequently, we collect the camera images and LiDAR point clouds, which are captured synchronously in a stationary setup, to serve as the primary input. Following data input, we obtain coarse calibrated intrinsic through either targetless methods leveraging SuperGlue [35] correspondences for perspective-n-point (PnP) estimation, or traditional checkerboards [5]/AprilTag-based [6] calibration approaches. For extrinsic calibration, we leverage CAD values from assembly drawings, which benefit from the high precision achievable with modern

[1] https://github.com/sys111111/Joint_Direct_LiDAR_Camera_Calibration

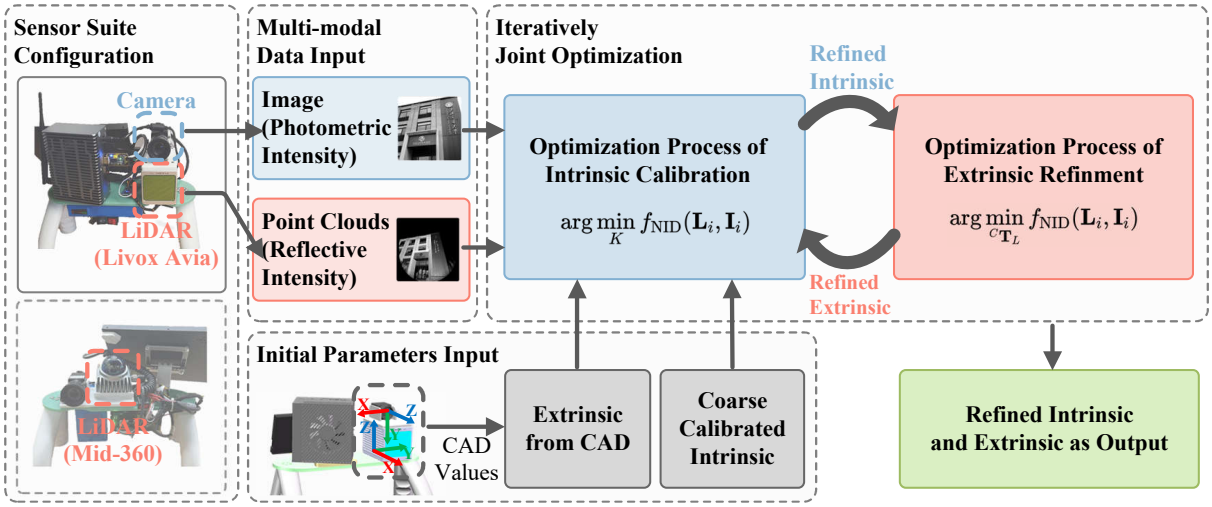


Fig. 2: Overview of our LiDAR-Camera joint calibration method using NID-based optimization. We employ distinct sensor suites for the acquisition of images and point clouds as data inputs. Subsequently, we utilize the initial intrinsic and extrinsic to perform intrinsic refinement through NID, followed by the joint refinement procedure that simultaneously optimizes extrinsic and intrinsic calibration parameters.

manufacturing techniques, such as 3D printing, where inter-component accuracy can approach sub-millimeter levels. Ultimately, we iteratively refine both intrinsic and extrinsic using the NID metric.

A. Coarse Intrinsic Calibration

For intrinsic initialization, we may start with either the coarse calibration method or use target-based methods to provide initial values. Similar to the method in [10], we can obtain correspondences between LiDAR and camera images using SuperGlue [35] and SuperPoint [36] to establish the alignment between LiDAR intensity values and image intensity values across the overlapping field of view.

Given a set of 3D LiDAR points ${}^L\boldsymbol{p}_j \in \mathbf{P}_i$ and their corresponding 2D image coordinates $\boldsymbol{q}_j$ obtained from established matches, we formulate the intrinsic optimization as a reprojection error minimization problem. To ensure the reliability of the correspondences, we implement a rigorous outlier rejection strategy and configure the matching thresholds to critical values that minimize false matches. The reprojection error is calculated as:

$$\boldsymbol{K}^*, \boldsymbol{D}^* = \arg\min_{\boldsymbol{K},\boldsymbol{D}} \sum_{j=1}^{M} \left\| \pi \left({}^C\mathbf{T}_L \, {}^L\boldsymbol{p}_j \right) - \boldsymbol{q}_j \right\|^2 \tag{1}$$

where the camera's intrinsic parameters are represented by the intrinsic matrix $\boldsymbol{K}$ and the distortion coefficients $\boldsymbol{D}$. For perspective cameras, $\boldsymbol{K}$ contains f_x, f_y, c_x and c_y, where f_x, f_y denote the focal lengths and c_x, c_y denote the principal point coordinates. The distortion coefficients $\boldsymbol{D}$ contain the radial distortion parameters and the tangential distortion parameters. π denotes the projection function mapping 3D points to 2D image coordinates using $\boldsymbol{K}$ and $\boldsymbol{D}$. ${}^C\mathbf{T}_L$ is the 6 DoF extrinsic between LiDAR and camera, representing the extrinsic of the LiDAR frame with respect to the camera coordinate frame. This optimization yields coarse intrinsic for further refinement in our subsequent NID-based optimization process.

B. NID Error Metric

Traditional geometric error metrics frequently prove inadequate for capturing the complex intensity relationships between camera and LiDAR modalities. To address this limitation, we leverage information theory by adopting the Normalized Information Distance (NID) between LiDAR reflectivity values and their corresponding camera photometric intensities as our optimization criterion. Compared to other mutual information metrics, such as Mutual Information (MI), NID exhibits superior environmental adaptability [37].

For the computation of NID, we extract the paired intensity values from both modalities. These values are binned to construct probability distributions: $P(\mathbf{L}_i)$ for LiDAR intensities, $P(\mathbf{I}_i)$ for image intensities, and $P(\mathbf{L}_i, \mathbf{I}_i)$ for their joint distribution. The entropies of the distributions $\mathrm{H}(\mathbf{L}_i)$, $\mathrm{H}(\mathbf{I}_i)$ and $\mathrm{H}(\mathbf{L}_i, \mathbf{I}_i)$ are computed as:

$$\mathrm{H}(X) = -\sum_{x \in X} \mathrm{P}(x) \log \mathrm{P}(x) \tag{2}$$

where x represents intensity values within each bin of the histogram. The NID between modalities is then calculated as:

$$f_{\text{NID}}(\mathbf{L}_i, \mathbf{I}_i) = 2 - \frac{\mathrm{H}(\mathbf{L}_i) + \mathrm{H}(\mathbf{I}_i)}{\mathrm{H}(\mathbf{L}_i, \mathbf{I}_i)} \tag{3}$$

We employ the NID metric for both intrinsic and extrinsic optimization to construct the loss function, achieving precise cross-modal alignment that consistently outperforms conventional calibration techniques. As calibration parameters approach their optimal values, the NID between projected

LiDAR and image intensities minimizes, indicating improved registration quality.

C. Joint Optimization for intrinsic and extrinsic

1) Intrinsic Refinement: Target-based calibration methods exhibit inherent limitations in achieving sub-pixel accuracy, with precision degrading non-uniformly across varying depths in three-dimensional space—particularly manifesting increased measurement error at extended distances. Additionally, feature extraction methods frequently fail to provide sufficient observation points for calibration. In contrast, LiDAR-based measurements maintain consistent accuracy irrespective of depth. Leveraging the abundant observation points and extended-range measurement capabilities of LiDAR, we extend the Normalized Information Distance (NID) optimization method to simultaneously refine camera intrinsic.

For a set of LiDAR point clouds ${}^{L}\boldsymbol{p}_j \in \mathbf{P}_i$, we project them onto the image plane using $\boldsymbol{q}_j = \pi\left({}^{C}\mathbf{T}_L {}^{L}\boldsymbol{p}_j\right)$. The camera intrinsic matrix $\boldsymbol{K}$ of π is optimized by minimizing the NID between LiDAR intensities and corresponding image intensities:

$$\boldsymbol{K}^* = \arg\min_{\boldsymbol{K}} f_{\text{NID}}\left(\mathbf{L}_i, \mathbf{I}_i\left(\boldsymbol{K}\left({}^{C}\mathbf{T}_L {}^{L}\boldsymbol{p}_j\right)\right)\right) \tag{4}$$

We apply the Broyden Fletcher Goldfarb Shanno (BFGS) algorithm through Ceres Solver to optimize the photometric and reflection intensity relationships. Additionally, the BFGS optimizer eliminates the need to compute the Jacobian matrix of our non-linear objective function, making the optimization process more efficient.

Notably, the changes in the camera projection model alter the perspective relationship between point clouds and the image plane during intrinsic parameter optimization. This transformation induces changes in the occlusion relationships between 3D point clouds and the 2D image plane, particularly at depth-discontinuous boundaries, resulting in the multi-valued and zero-valued mapping phenomenon identified by Yuan et al. [11]. To address this challenge, we applied view-based hidden point removal to systematically filter non-visible LiDAR point clouds from the current camera viewpoint during the optimization process, ensuring that only visible points contribute to the NID computation.

2) Extrinsic Refinement: Following the refinement of intrinsic, we apply an identical NID-based optimization to the extrinsic. Our method implements an iterative refinement that optimizes the extrinsic ${}^{C}\mathbf{T}_L$ by minimizing:

$${}^{C}\mathbf{T}_L{}^{*} = \arg\min_{{}^{C}\mathbf{T}_L} f_{\text{NID}}\left(\mathbf{L}_i, \mathbf{I}_i\left(\boldsymbol{K}\left({}^{C}\mathbf{T}_L {}^{L}\boldsymbol{p}_j\right)\right)\right) \tag{5}$$

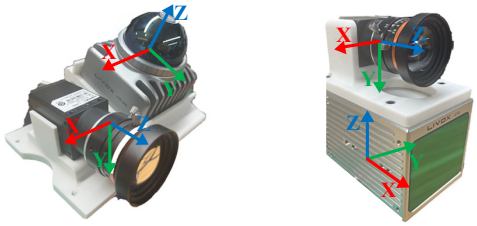
Similar to intrinsic optimization, we execute viewpoint updates upon parameter convergence within minimal iteration cycles, ensuring proper management of depth-discontinuous boundaries. This algorithm enables intrinsic calibration accuracy beyond the sub-pixel error to address limitations inherent in traditional methods. The precision of extrinsic calibration fundamentally depends on accurate intrinsic; hence, we employ an iterative optimization strategy between intrinsic and extrinsic until the convergence of both intrinsic and extrinsic is achieved.

IV. Experiments

We present comprehensive experiments to evaluate the accuracy and robustness of our proposed method across various sensor configurations. We first assess the accuracy of intrinsic parameter optimization, followed by evaluating the extrinsic calibration results. Finally, we demonstrate the qualitative performance through LiDAR point clouds colorization results.

A. Experimental Setup

In our experiments, we utilized both public and self-integrated datasets. For proprietary data collection, we employed two self-integrated sensor suites as shown in Fig. 3: Sensor Suite 1 (Mid-360 LiDAR with pinhole camera) and Sensor Suite 2 (Livox Avia LiDAR with pinhole camera). Data collection spanned diverse environments with varying illumination conditions and structural complexities in Shanghai Jiao Tong University campus. Indoor datasets (Data 1 and Data 2) were collected using Sensor Suite 1 and Sensor Suite 2, respectively, while outdoor datasets (Data 3 and Data 4) were acquired following the same corresponding sensor suites. To evaluate the accuracy of our method in complex, large-scale outdoor scenarios, we collected Data 5 using Sensor Suite 2, comprising multiple sequences of a single large-scale scene from diverse viewpoints. All experiments were conducted on a workstation equipped with an Intel i7 processor and 32GB RAM.



(1) Mid-360 LiDAR with pinhole camera **(2) Livox Avia LiDAR with pinhole camera**

Fig. 3: Sensor suites for calibration experiments.

B. Intrinsic Calibration Evaluation

We conduct a comprehensive evaluation of our intrinsic calibration method against the edge-based targetless joint calibration method proposed by [31] using their public dataset comprising four diverse scenes collected at The University of Hong Kong campus. As illustrated in the methodology section, the NID metric serves as our evaluation criterion, where lower values indicate more accurate intrinsic. As shown in Table I, our method demonstrates better intrinsic calibration accuracy and robustness in different scenes, achieving lower NID cost (0.960 and 0.990).

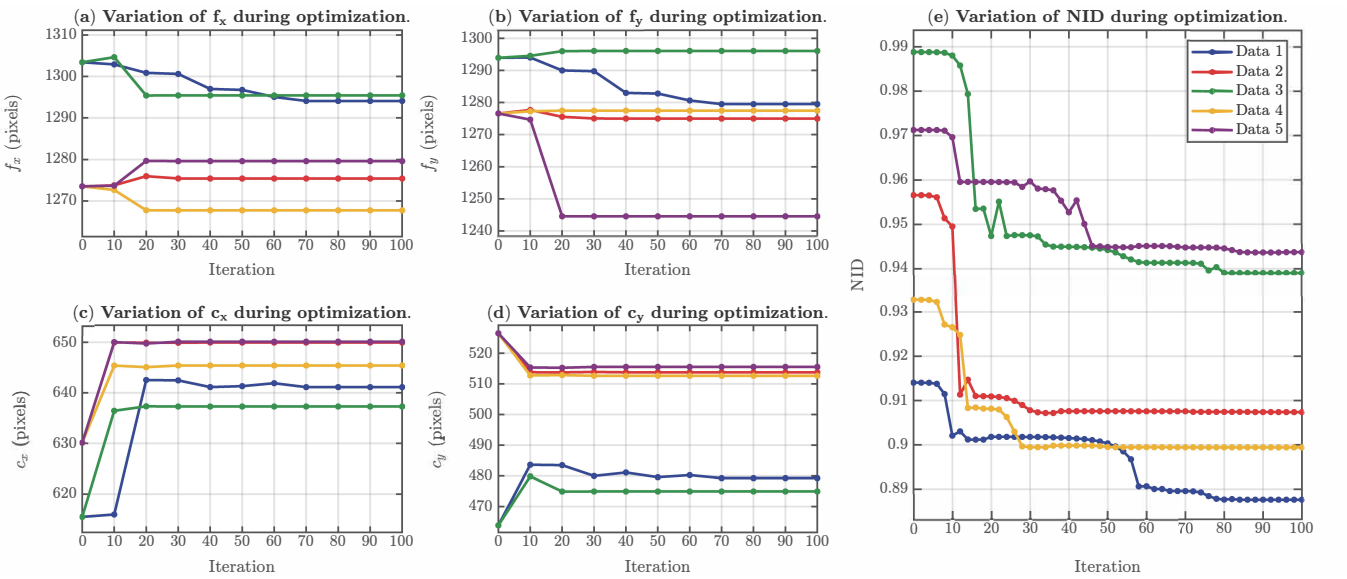


Fig. 4: Evolution of five datasets camera intrinsic and NID value during the optimization process.

TABLE I: Intrinsic calibration evaluation using the public datasets of HKU campus

Dataset	Our Method NID↓	Joint Edge-based Method [31] NID↓
Scene 1	**0.954**	0.988
Scene 2	**0.961**	0.991
Scene 3	**0.957**	0.989
Scene 4	**0.969**	0.992
Average	**0.960**	0.990

TABLE II: Intrinsic calibration evaluation across different sensor suites

Dataset	Our Method		Target-based Method [7]	
	Projection Error (Pixel)↓	NID ↓	Projection Error (Pixel)↓	NID ↓
Data 1	0.180	**0.885**	**0.179**	0.920
Data 2	**0.178**	**0.874**	0.180	0.903
Data 3	**0.173**	**0.881**	0.174	0.918
Data 4	**0.180**	**0.899**	**0.180**	0.931
Data 5	**0.174**	**0.909**	0.176	0.942
Average	**0.177**	**0.890**	0.178	0.923

We also evaluate our method against the target-based intrinsic calibration method Kalibr, proposed by [7]. We employ the projection error and NID metric to assess intrinsic accuracy. The initial intrinsic for both experimental setups is obtained through the target-based calibration, ensuring a fair comparison baseline. Table II presents a comparative analysis across five diverse self-collected datasets, where the projection error results are evaluated using checkerboards for intrinsic calibration. The results demonstrate that our targetless calibration method achieves comparable projection error (0.177 and 0.178 on average), confirming beyond sub-pixel high accuracy surpassing target-based methods. Meanwhile, our method consistently delivers lower NID values across all datasets (0.890 and 0.923 on average).

As shown in Fig. 4, our optimization method demonstrates stable convergence of intrinsic over iterations. For this experiment, initial intrinsic is derived from feature correspondences established through SuperGlue. The intrinsic parameters, f_x, f_y, c_x and c_y gradually stabilize while the NID value remains consistent, confirming that our method effectively refines camera intrinsic with improving the overall registration quality between LiDAR and camera.

C. Extrinsic Calibration Evaluation

We evaluated our extrinsic calibration method against several state-of-the-art methods: an NID-based method [10], an edge-based method [11] and a joint edge-based calibration method [31]. For each method, we assessed both the adherence to CAD specifications and the quality of registration as measured by the NID metric. For fairness, the intrinsic employed in the extrinsic calibration evaluation were derived using the target-based approach.

Table III presents the deviation from CAD specifications in translation (meters) as well as the corresponding NID values for each method across the five self-collected datasets. Our method exhibits superior CAD accuracy compared to existing methods, achieving an average Euclidean error of 0.049 m, which is a substantial reduction in deviation from physical specifications. Notably, while NID-based method [10] achieves similar NID average values, they frequently produce extrinsic that substantially diverge from CAD values.

A critical observation from our experiments is that current methods often achieve visual alignment between LiDAR point clouds and camera images while producing extrinsic that significantly deviate from CAD values. Errors in focal

TABLE III: Extrinsic calibration evaluation across different sensor suites and methods

Dataset	Our Method		Targetless Method [10]		Edge-based Method [11]		Joint Edge-based Method [31]	
	Euclidean Error (m) ↓	NID ↓	Euclidean Error (m) ↓	NID ↓	Euclidean Error (m) ↓	NID ↓	Euclidean Error (m) ↓	NID ↓
Data 1	**0.025**	0.884	0.073	**0.882**	0.315	0.924	0.342	0.978
Data 2	**0.011**	**0.876**	0.021	0.878	0.267	0.921	0.442	0.986
Data 3	**0.079**	**0.877**	0.164	0.881	0.295	0.919	0.409	0.975
Data 4	**0.056**	**0.896**	0.060	0.897	0.230	0.918	0.342	0.922
Data 5	**0.075**	**0.897**	0.109	**0.897**	0.344	0.932	0.498	0.987
Average	**0.049**	**0.886**	0.085	0.887	0.290	0.923	0.406	0.970

length parameters f_x, f_y often manifest as errors in the depth component of the translation vector, while errors in principal point parameters c_x, c_y affect the other two translation components. Our method mitigates this issue by starting with initial CAD values and first optimizing intrinsic before proceeding to extrinsic refinement, ensuring that the optimized extrinsic remain within a reasonable range of their true CAD values.

D. Qualitative Evaluation

To qualitatively evaluate the performance of our method, we use the visualization tool modified from Koide et al. [10] to visualize the colorized point. Figure 5 shows representative results across various sensor suites.

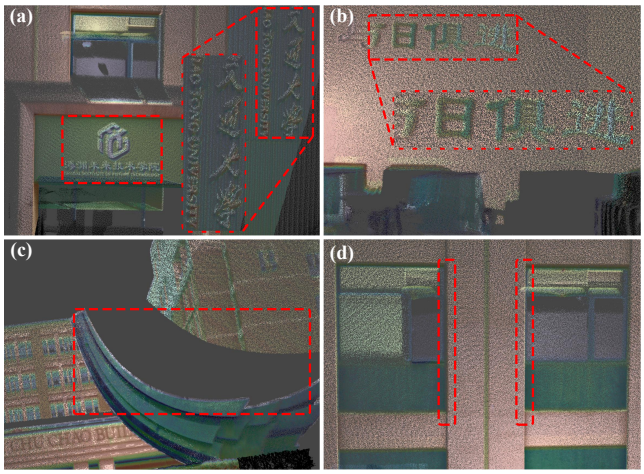


Fig. 5: Qualitative visualization of LiDAR-camera calibration results from four different scenes (a)-(d). The image shows blended rendering of camera photometric intensity and LiDAR reflectivity intensity, demonstrating precise alignment through our calibration method.

Notably, existing methods [10, 11] also achieve visually satisfactory point cloud colorization, demonstrating that various combinations of intrinsic and extrinsic yield seemingly correct alignments. It is worth noting that our method distinguishes itself as the only method that simultaneously delivers high visual alignment quality while maintaining strict adherence to CAD specifications. Additional experiments indicate our method can maintain satisfactory colorization even under slight perturbations to the extrinsic parameters.

E. Computational Efficiency

Our proposed method demonstrates superior computational efficiency compared to current mixture calibration methods that separately perform intrinsic [7] calibration using checkerboards and edge-based extrinsic calibrations [11], as well as the joint edge-based method [31]. Table IV compares the processing times required for each method. Our complete calibration process achieves an average processing time of approximately 262 seconds on our workstation, substantially decreasing the computational time. This improvement in computational efficiency is primarily due to our direct NID-based 3D-2D matching calibration method, which circumvents the computationally intensive feature extraction steps inherent to separate target-based methods [7, 11] and joint edge-based methods [31]. Additionally, the implementation of an alternating optimization method integrated with the efficient BFGS algorithm within Ceres Solver further expedites convergence, reinforcing the overall computational advantage of our proposed calibration method.

TABLE IV: Total processing time comparison (s)

Dataset	Our Method	Mixture Method [7, 11]	Joint Edge-based [31]
Data 1	**306**	482	509
Data 2	**214**	398	452
Data 3	**265**	422	417
Average	**262**	434	459

V. Conclusion

In this paper, we address the critical challenge of sub-pixel intrinsic calibration accuracy and its impact on extrinsic precision in LiDAR-camera calibration. We present the direct targetless calibration method employing normalized information distance (NID) as the optimization criterion within two stages: first, refining camera intrinsic beyond conventional sub-pixel limitations using LiDAR's precise measurements, then optimizing extrinsic while maintaining CAD specification consistency. Our approach effectively captures parameter interdependence while preventing physically implausible

solutions. Experiments demonstrate our method's superior intrinsic and extrinsic calibration accuracy.

Future work will extend our approach to additional sensor modalities, particularly fisheye cameras with pronounced distortion, and develop more efficient optimization strategies using advanced gradient methods to mitigate potential non-convergence issues, which may stem from inaccurate coarse intrinsic estimates or limited scene texture features.

REFERENCES

[1] C. Zheng, W. Xu, Z. Zou, T. Hua, C. Yuan, D. He, B. Zhou, Z. Liu, J. Lin, F. Zhu *et al.*, "Fast-livo2: Fast, direct lidar-inertial-visual odometry," *IEEE Transactions on Robotics*, 2024.

[2] S. Hong, C. Zheng, Y. Shen, C. Li, F. Zhang, T. Qin, and S. Shen, "Gs-livo: Real-time lidar, inertial, and visual multi-sensor fused odometry with gaussian mapping," *arXiv preprint arXiv:2501.08672*, 2025.

[3] J. Yu, Y. Jiang, Z. Wang, Z. Cao, and T. Huang, "Unitbox: An advanced object detection network," in *Proceedings of the 24th ACM international conference on Multimedia*, 2016, pp. 516–520.

[4] M. Ye, S. Xu, and T. Cao, "Hvnet: Hybrid voxel network for lidar based 3d object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[5] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[6] W. Gao, J. Lin, F. Zhang, and S. Shen, "A screen-based method for automated camera intrinsic calibration on production lines," in *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2019, pp. 392–398.

[7] L. Oth, P. Furgale, L. Kneip, and R. Siegwart, "Rolling shutter camera calibration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1360–1367.

[8] M. Smith and E. Cope, "The effects of temperature variation on single-lens-reflex digital camera calibration parameters," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, no. Part 5, 2010.

[9] F. Devernay and O. Faugeras, "Straight lines have to be straight," *Machine Vision and Applications*, vol. 13, no. 1, pp. 14–24, Aug. 2001.

[10] K. Koide, S. Oishi, M. Yokozuka, and A. Banno, "General, single-shot, target-less, and automatic lidar-camera extrinsic calibration toolbox," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 11 301–11 307.

[11] C. Yuan, X. Liu, X. Hong, and F. Zhang, "Pixel-level extrinsic self calibration of high resolution LiDAR and camera in targetless environments," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7517–7524, Oct. 2021.

[12] E.-s. Kim and S.-Y. Park, "Extrinsic calibration between camera and lidar sensors by matching multiple 3d planes," *Sensors*, vol. 20, no. 1, p. 52, 2019.

[13] J. L. Schönberger and J.-M. Frahm, "Structure-from-motion revisited," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 4104–4113.

[14] S. J. Maybank and O. D. Faugeras, "A theory of self-calibration of a moving camera," *International Journal of Computer Vision*, vol. 8, pp. 123–151, 1992. [Online]. Available: https://api.semanticscholar.org/CorpusID:20984115

[15] Q.-T. Luong and O. D. Faugeras, "Self-calibration of a moving camera from pointcorrespondences and fundamental matrices," *Int. J. Comput. Vision*, vol. 22, no. 3, p. 261–289, Mar. 1997. [Online]. Available: https://doi.org/10.1023/A:1007982716991

[16] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A toolbox for easily calibrating omnidirectional cameras," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2006, pp. 5695–5701.

[17] K. Kanatani, "Calibration of ultrawide fisheye lens cameras by eigenvalue minimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 4, pp. 813–822, 2012.

[18] J. Kannala and S. S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 8, pp. 1335–1340, 2006.

[19] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Sep. 2004, pp. 2301–2306.

[20] C. Fang, S. Ding, Z. Dong, H. Li, S. Zhu, and P. Tan, "Single-shot is enough: Panoramic infrastructure based calibration of multiple cameras and 3d LiDARs," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Sep. 2021, pp. 8890–8897.

[21] J. Beltran, C. Guindel, A. de Escalera la, and F. Garcia, "Automatic extrinsic calibration method for LiDAR and camera sensor setups," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, Mar. 2022.

[22] L. Zhou, Z. Li, and M. Kaess, "Automatic extrinsic calibration of a camera and a 3d LiDAR using line and plane correspondences," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Oct. 2018, pp. 5562–5569.

[23] X. Gong, Y. Lin, and J. Liu, "3d lidar-camera extrinsic calibration using an arbitrary trihedron," *Sensors*, vol. 13, no. 2, pp. 1902–1918, 2013.

[24] Q. Liao, Z. Chen, Y. Liu, Z. Wang, and M. Liu, "Extrinsic calibration of lidar and camera with polygon," in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2018, pp. 200–205.

[25] X. Zhang, S. Zhu, S. Guo, J. Li, and H. Liu, "Line-based automatic extrinsic calibration of LiDAR and camera," in *IEEE International Conference on Robotics and Automation*. IEEE, May 2021, pp. 9347–9353.

[26] X. Liu, C. Yuan, and F. Zhang, "Targetless extrinsic calibration of multiple small FoV LiDARs and cameras using adaptive voxelization," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, May 2022.

[27] Q. Ye, L. Shu, and W. Zhang, "Extrinsic calibration of a monocular camera and a single line scanning lidar," in *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE, 2019, pp. 1047–1054.

[28] J. Levinson and S. Thrun, "Automatic online calibration of cameras and lasers." in *Robotics: science and systems*, vol. 2, no. 7. Citeseer, 2013.

[29] J. Kang and N. L. Doh, "Automatic targetless camera–lidar calibration by aligning edge with gaussian mixture model," *Journal of Field Robotics*, vol. 37, no. 1, pp. 158–179, 2020.

[30] G. Yan, F. He, C. Shi, P. Wei, X. Cai, and Y. Li, "Joint camera intrinsic and lidar-camera extrinsic calibration," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 11 446–11 452.

[31] L. Li, H. Li, X. Liu, D. He, Z. Miao, F. Kong, R. Li, Z. Liu, and F. Zhang, "Joint intrinsic and extrinsic lidar-camera calibration in targetless environments using plane-constrained bundle adjustment," *arXiv preprint arXiv:2308.12629*, 2023.

[32] A. Stewart, "Localisation using the appearance of prior structure," Ph.D. dissertation, University of Oxford, 2014.

[33] J. Jeong, Y. Cho, and A. Kim, "The road is enough! extrinsic calibration of non-overlapping stereo camera and lidar using road information," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2831–2838, 2019.

[34] G. Pascoe, W. Maddern, M. Tanner, P. Pinies, and P. Newman, "Nid-slam: Robust monocular slam using normalised information distance," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[35] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperGlue: Learning feature matching with graph neural networks," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Jun. 2020, pp. 4938–4947.

[36] D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperPoint: Self-supervised interest point detection and description," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. IEEE, Jun. 2018, pp. 224–236.

[37] B. Patel, *Visual localization for UAVs in outdoor GPS-denied environments*. University of Toronto (Canada), 2019.