ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-G-2025
ISPRS Geospatial Week 2025 "Photogrammetry & Remote Sensing for a Better Tomorrow…", 6–11 April 2025, Dubai, UAE

MapCalib: Toward Large-Scale Automatic Calibration for Roadside LiDAR Using High-Definition Map

Cong Duan[1], Haomiao Bian[2], Yuan Guo[3], Bijun Li[1], Maosheng Yan[1], Jian Zhou[1]

[1] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China-(duancong, lee, yanms sgg, jianzhou)@whu.edu.cn

[2] School of remote sensing, Wuhan University, China-haomiaobian@whu.edu.cn

[3] School of Resources and Environmental Engineering, Wuhan University of Technology, China-yuang@whut.edu.cn

Cong Duan and Haomiao Bian contributed equally to this work

Keywords: Roadside LiDAR Calibration, HD map, Virtual Map Projector, Semantic Universal Spatial Context.

Abstract

Roadside LiDAR sensors are critical components in automated driving systems and mobile mapping systems. These sensors, typically deployed along roadsides to provide continuous data for perception tasks, require precise calibration to ensure the safety and performance of intelligent connected vehicles. However, large-scale deployment presents new challenges, including low sensor overlap and variability in sensor types, complicating the calibration process. To address these challenges, this study introduces MapCalib, an innovative method for the automatic calibration of roadside LiDAR systems using High-Definition (HD) map. MapCalib improves calibration efficiency by eliminating the need for specific calibration targets, which simplifies the process and increases safety when managing large-scale roadside LiDAR installations. The method begins with the development of a virtual map projector, which establishes a mapping from the HD map to the LiDAR data, minimizing representation disparities. Next, a Semantic Universal Spatial Context (SUSC) descriptor is proposed to efficiently localize the LiDAR sensor positions within the HD map. Finally, through feature retrieval and iterative optimization, the method calibrates sensor parameters, such as orientation and position. The proposed calibration framework is validated through simulated, public, and self-collected datasets, demonstrating its ability to automatically calibrate multiple LiDAR sensors with high accuracy. Compared to existing calibration methods, MapCalib achieves a notable improvement of 39.9% in Relative Rotation Error (RRE) and 39.3% in Relative Translation Error (RTE).

1. Introduction

With the advancement of 5G technology (Hakak et al., 2023) and high-definition (HD) map (Liu et al., 2020), vehicle-road cooperative systems have become the predominant approach in modern autonomous driving systems. By integrating multi-perspective observations from vehicle and roadside sensors, this collaborative perception approach addresses challenges such as limited sensor ranges and blind spots caused by environmental occlusions (Kulawiak, 2024). However, the use of different coordinate reference systems by vehicle and roadside sensors presents substantial challenges for effective collaboration across various observation nodes (Kim et al., 2014). Addressing these challenges requires joint calibration of sensors based on a unified coordinate reference system, which is crucial for effective multi-source data integration in vehicle-road collaborative sensing.

Among various roadside sensing modalities, Light Detection and Ranging (LiDAR) has emerged as a pivotal component due to its high-resolution 3D measurement capabilities, playing crucial roles in critical applications ranging from autonomous driving (Li and Ibanez-Guzman, 2020) to simultaneous localization and mapping (SLAM) (Khan et al., 2021). To acquire an accurate georeferenced roadside LiDAR point cloud, it is essential to precisely estimate the parameters of the roadside LiDAR. Various LiDAR calibration methods have been developed over the past few decades. Traditional calibration methods typically rely on specific targets such as Apriltags (Xie et al., 2018), checkerboard (Lee et al., 2020) and spheres (Zhang et al., 2024). However, these methods necessitate road closures during the calibration process to ensure operator safety. Additionally, the extensive and costly manual operations render them inefficient for large-scale deployment. Consequently, achieving large-scale automated calibration for roadside LiDAR has become a pressing challenge. With advancements in sensor perception capabilities, target-free calibration in outdoor scenarios has progressively become the mainstream approach (Zhang et al., 2022; Herau et al., 2024). These methods generally require a significant overlap between sensors and assume identical sensor types, solving the calibration problem by matching features widely found in natural environments, such as linear features (Zhou et al., 2018), planar features (Jiao et al., 2019), and semantic edge features (Liao et al., 2023). However, large-scale roadside LiDAR calibration is challenging due to the wide distribution, fixed positioning, and narrow field of view of roadside LiDAR, resulting in low overlap and variations between LiDAR types.

To address these challenges, we propose an innovative roadside LiDAR calibration framework using HD map. HD map not only provides detailed, full-coverage observational data of roads but also offers precise georeferencing, making them an ideal medium for bridging gaps between sensor data. By incorporating HD map as auxiliary data, our approach effectively compensates for spatial gaps, physical discrepancies, and variations in sensor viewpoints. This integration enhances the robustness of calibration across diverse roadside LiDAR systems, bridging the gaps between different sensors, and ensuring consistent performance in large-scale deployment scenarios.

In this paper, we introduce MapCalib, a novel automatic calibration method for roadside LiDAR using HD map. To address representation discrepancies between LiDAR and HD map, we propose a virtual map projector that performs virtual mapping from the HD map to roadside LiDAR, effectively reducing data discrepancies. Additionally, the diversity in the types, numbers, fixed positions, and viewing angles of roadside LiDAR systems introduces challenges for large-scale calibration. To address this, we present a Semantic Universal Spatial Context (SUSC) descriptor, tailored to various LiDAR types, which accounts for installation methods and physical mechanisms, ensuring consistent data representation. Finally, through feature retrieval and iterative optimization, the extrinsic parameters of the roadside LiDAR are refined to ensure precise calibration.

2. Related Work

2.1 Roadside Sensor Calibration

Calibration aims to find the transformation between two sensor data, including relative rotation, and translation. This process is typically classified into two categories: target-based and target-less methods. Target-based calibration relies on specially designed targets that sensors can precisely track, offering a reference for calibration. In contrast, target-less methods derive calibration parameters directly from environmental data without the need for a physical target.

Target-based calibration methods are typically inefficient and may pose safety risks. As a result, target-less calibration methods have become the primary focus of current research. For instance, Zhang et al. (2020) accomplished calibration by utilizing retro-reflective materials and employing point cloud registration. VI-eye (He et al., 2021) is a real-time system for registering vehicle infrastructure point clouds, achieving centimeter-level accuracy through detecting a set of key semantic objects. TrajMatch (Ren et al., 2023) is an automatic roadside LiDAR calibration system that calibrates sensors based on detection and tracking task results. These calibration methods typically focus on the relative relationship between sensors, overlooking the connection to the absolute coordinate system. This limitation hinders the ability to achieve sensor calibration across different scenes. To address this issue, Zhang et al. (2022) proposed a roadside millimeter-wave radar calibration method that leverages real-time traffic data and HD map. This method clusters vehicle trajectories, fits lane centerlines using polynomials, performs uniform sampling, and solves calibration parameters via nonlinear optimization, enabling accurate calibration without road closures. However, this method relies on manually defined regions of interest, limiting its level of automation. Zhao et al. (2024) employed positioning and perception data from autonomous vehicles to calibrate roadside sensors. However, the proposed calibration framework depends on obtaining object-level trajectory data, limiting its applicability for calibrating raw perception data.

To address these issues, we propose MapCalib, a novel roadside LiDAR calibration method that leverages HD map as auxiliary input, enabling large-scale automatic calibration of roadside LiDAR systems using only raw data.

2.2 HD Map

HD map is crucial for highly automated driving systems (Xiao et al., 2024). This map provides comprehensive data about the vehicle's environment, mitigating sensor limitations by compensating for occluded areas or regions beyond the sensor's range. Moreover, in cases of sensor uncertainty, the HD map serves as a reliable reference, enabling accurate interpretation and decision-making across a wide range of driving scenarios.

Typically, HD map consists of road networks, lane information, localization features, and traffic infrastructure data. The road network primarily includes geometric and attribute-based information, such as road type, classification, and width. Lane information focuses on lane markings, offering detailed specifications on direction, number, and speed limits in each lane. Localization features consist of reference points used by autonomous vehicles for localization, specifying their location, type, texture, and shape. The signal layer incorporates geometric and semantic data regarding traffic signs, lights, and road markings, detailing their type, height, and other attributes. The road network in HD map aids autonomous vehicles with global navigation and path planning, while lane information enables fine-grained, lane-level path planning in automated driving. Localization features and signal data contribute to environmental perception and precise localization for autonomous vehicles.

Building upon the work of Wong et al. (2021), Srinara et al. (2022), and Chiang et al. (2023) in integrating point cloud maps with LiDAR-IMU calibration, this study investigates the processing and analysis of point cloud data, which plays a critical role in the HD map. By aligning sensor data with the HD map, precise sensor calibration is attained.

3. Methodology

The proposed MapCalib framework is designed to take full advantage of the data characteristics of roadside LiDAR and the absolute pose reference provided by HD map, addressing the challenge of large-scale automatic roadside LiDAR calibration. The workflow of this method is depicted in Figure 1. Initially, a virtual map projector is designed to facilitate virtual mapping from the HD map to the roadside LiDAR by incorporating the road's geometric layout and spatial characteristics. Subsequently, a Semantic Universal Spatial Context (SUSC) descriptor is developed. Ultimately, through feature retrieval and iterative optimization, the extrinsic parameters of the roadside LiDAR are refined to achieve accurate calibration.

3.1 Virtual Map Projector

The HD map for autonomous driving represents a novel integration of map science and innovations within the automotive industry, attracting significant attention across academic research, government regulation, and industrial applications (Guo et al., 2024). However, significant differences in density and coverage often exist between existing HD map and roadside LiDAR, complicating the direct unification of their coordinate systems. To address this, a virtual map projector is proposed to process the HD map, achieving virtual mapping that registers with roadside LiDAR.

Virtual Viewpoint Generation: To construct the virtual map projector, it is essential to determine the pose of each virtual viewpoint in the HD map, which is critical for ensuring the accuracy and robustness of the projector. This study proposes a road boundary-based search strategy, aligned with the practical deployment of roadside LiDAR. First, road boundary information is extracted from the HD map, with each boundary represented by a series of continuous vectors, as shown in Equation

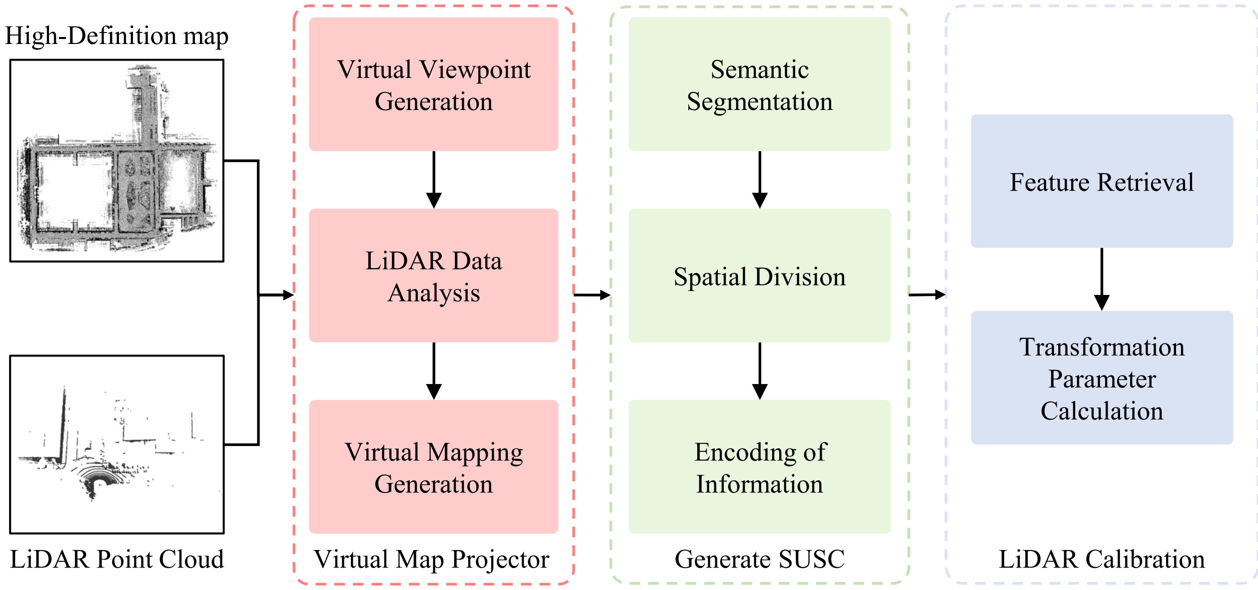

Figure 1. Architecture overview. The proposed MapCalib comprises three main components: 1) Virtual Map Projector, 2) Generation of Semantic Universal Spatial Context (SUSC) Descriptor, and 3) Finally Calibration. The Virtual Map Projector constructs a virtual mapping from HD map to roadside LiDAR, while the SUSC Descriptor is designed to locate the LiDAR position within the HD map. Finally, the feature retrieval and iterative optimization are used to achieve precise calibration of the roadside LiDAR.

(1)

$$\begin{cases} R_i = \{V_{i,j}\}_{j=1}^{n_i-1} \\ V_{i,j} = P_{i,j+1} - P_{i,j} \end{cases} \quad (1)$$

where R_i denotes the i_{th} road, $V_{i,j}$ denotes the j_{th} vector composing the i_{th} road, n_i denotes the number of coordinate points on the i_{th} road, and $P_{i,j}$ denotes the j_{th} coordinate point on the i_{th} road.

To generate the virtual viewpoints, linear interpolation is used between the start and end points of each vector. These virtual viewpoints are distributed at regular intervals along the road boundary, as described in Equation (3), ensuring that the viewpoints are spaced to cover the desired resolution.

$$m = \left\lceil \frac{P_{i,j}^{\text{end}} - P_{i,j}^{\text{start}}}{\text{interpert}}, \right\rceil \quad (2)$$

$$\begin{aligned} P_i(k,j) = &\left(1 - \frac{k}{m-1}\right) P_{i,j}^{\text{start}} \\ &+ \frac{k}{m-1} P_{i,j}^{\text{end}}, \quad k = 0, 1, 2, \ldots, m-1 \end{aligned} \quad (3)$$

where $P_i(k)$ denotes the k-th virtual viewpoint generated on the i-th road. m denotes the number of virtual viewpoint desired to be generated. $P_{i,j}^{start}$ denotes the start point of the j-th vector in the i-th road, and $P_{i,j}^{end}$ denotes the end point of the j-th vector in the i-th road. $interpert$ denotes the interpolation interval. Note that the Z-axis is uniformly set to the same height because the roadside LiDAR installation height is relatively fixed. For the attitude information, considering the rotational invariance of the feature descriptors, we set the yaw angle, roll angle and pitch angle to zero.

Roadside LiDAR Data Analysis: After obtaining the position, in order to simulate the real scanning as accurately as possible, the parameters are obtained by parsing the actual roadside LiDAR data. The roadside LiDAR data are first represented as $P_L = P_{L1}, P_{L2}, \cdots P_{Ln}$, with each point $P_{Li}(P_{Li}^x, P_{Li}^y, P_{Li}^z)$, for which the distance D_i, horizontal angle H_i, and vertical angle V_i are calculated, as shown in Equation (4)

$$\begin{cases} D_i = ||P_{Li}|| \\ V_i = \arctan\left(\frac{P_{Li}^z}{\sqrt{{P_{Li}^x}^2 + {P_{Li}^y}^2}}\right) \\ H_i = \arctan\left(\frac{P_{Li}^x}{P_{Li}^y}\right) \end{cases} \quad (4)$$

This formulation allows us to generate parameters that reflect the actual roadside LiDAR deployment. Two types of LiDAR systems are considered: mechanical LiDAR and semi-solid-state LiDAR.

- **Mechanical LiDAR:** This type operates by rotating the optical structure with a motor for full 360-degree scanning. Key parameters include the number of LiDAR lines, the measurement range, vertical field of view, and angular resolution. We simulate the number of vertical angles based on the real scanning data, use the farthest distance as the measurement range, and derive the angular resolution and vertical field of view from the data.
- **Semi-solid LiDAR:** This type alters the laser direction using moving mirrors. It employs vibrating mirrors and prism technology, facilitating dynamic adjustment of the region of interest (ROI). For semi-solid-state LiDAR, the ROI and non-ROI areas are identified based on vertical angular gaps, with distinct angular resolutions assigned accordingly.

Virtual Mapping Generation: Once the position and parameters are determined, the virtual map projector is generated within the HD map. Initially, the HD map is cropped based on the position and measurement range. Subsequently, all LiDAR scan lines are generated using horizontal and vertical angles. Points within the cropped map are evaluated to determine whether they lie on a scan line, with only those on the lines being retained. For each scan line, the point nearest to the virtual viewpoint is retained, resulting in the final virtual mapping.

3.2 Generation of SUSC Descriptor

Existing point cloud feature descriptors usually use point height or point density information as features, the essence of which is to take advantage of the particular properties of different objects in the scene. However, there are a wide variety of roadside LiDARs, and low-level height and point density information is not sufficient to fully reflect the common features of multiple LiDARs. Recent studies have demonstrated the importance of semantic information in point cloud descriptions (Li et al., 2021). Therefore, this paper introduces semantic information and proposes a semantic unified feature descriptor (SUSC), as shown in Figure 2.

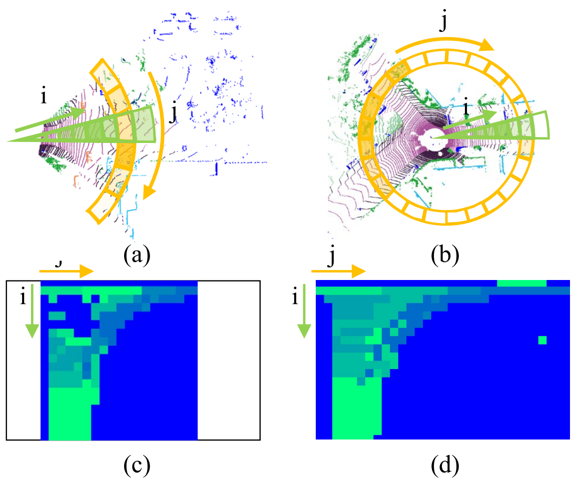

Figure 2. Semantic Universal Spatial Context descriptor creation. (a) and (b) represent the semantic segmentation results of two different LiDARs with different field-of-view angles. We divide the space around the vehicle into discrete regions on the x-y plane. The green area is a sector, and the yellow area is a ring. Their overlapping area is a bin. (c) and (d) represent the generated SUSC descriptors, where rows and columns represent the indices of the ring (i) and sector (j). Due to the different field-of-view angles of the two LiDARs, the number of columns in the generated descriptor is also different.

Specifically, in order to construct the SUSC descriptor, we choose RangeNet++, a CNN-based point cloud segmentation algorithm, for semantic segmentation processing of roadside LiDAR point clouds and their virtual mapping data from HD map. In our experiments, we used the pre-trained weights of RangeNet++ on the publicly available dataset SemanticKITTI for inference. To further improve the robustness of the feature descriptors, we filter the segmented point cloud to remove dynamic targets based on semantic labels and retain only static semantic objects. This step ensures that the descriptors are not interfered with dynamic objects in the scene, thus reflecting more accurately the key static structures in the scene. In addition, the filtering of static objects helps to improve the stability of features across different times and sensors.

Then, we divided the point cloud into bins at regular intervals along the azimuth and radial directions, followed by further segmentation along the vertical angle. After completing the division, $N_r \times N_h \times N_v$ bins can be obtained, where N_r, N_h, and N_v denote the number of bins along the radial, horizontal, and vertical angles, respectively. Notably, in order to improve the adaptability of the SUSC descriptor to different roadside LiDAR data, we design a flexible division mechanism that can dynamically adjust the corresponding division intervals according to the field-of-view angles and resolutions of different LiDARs, thus ensuring that the descriptor can effectively capture features in LiDAR systems with varied characteristics. As shown in (a) and (b) in Figure 2, the point cloud data in (a) has only 120-degree field-of-view angle, and thus is divided only within that sector, while the point cloud data in (b) represents 360-degree circumferential view data, and thus divides the space for the entire circular area.

After dividing all the points in the point cloud into bins, each bin is assigned a value based on the number of points in each semantic category and the elevation distribution within it. We first use the point density of each semantic category as the basis for environmental description and manually assign different weights to each category for normalization, emphasizing its importance. Additionally, elevation-based weighting is applied to enhance features nearer to the roadside LiDAR. The bin encoding function is given in Equation (5)

$$\Phi(P_{i,j,k}) = \sum_{l=1}^{N_{label}} \omega_l N(P_{i,j,k}, l) \times \frac{2^{k-1}}{\sum_{k=1}^{N_v} 2^{k-1}} \tag{5}$$

where $P_{i,j,k}$ denotes the set of points belonging to the $B_{i,j,k}(i \in [N_r], j \in [N_h], k \in [N_v])$, and $B_{i,j,k}$ denotes the k-th bin in the i-th ring and j-th sector. $N(P_{i,j,k}, l)$ denotes the number of points with semantic category l in the set of $P_{i,j,k}$. N_{label} denotes the number of semantic categories. ω_l denotes manually set weights for different semantic categories. After calculating each bin, we can represent the entire point cloud with a matrix of $N_r \times N_h$.

This coding function was selected because the use of elevation-weighted information yields a more accurate representation of the surrounding environment than relying solely on maximum height. Furthermore, due to the elevated mounting position of roadside LiDAR, data from higher elevations are more reliable and crucial in outdoor environments, whereas ground-level information tends to be redundant. Additionally, vehicles moving at lower heights may introduce noise, negatively impacting the results.

3.3 Feature Retrieval

After obtaining the SUSC descriptor for the roadside LiDAR point cloud and its corresponding virtual mapping, the distance between the LiDAR point cloud and each frame of the virtual mapping is calculated first. Considering that the virtual mapping field of view, constructed through the virtual reprojection model, is set to 360 degrees, while real roadside LiDAR, diverse in type, covers multiple scenarios—most of which have a field of view smaller than 360 degrees—the dimensions of the feature descriptors between the two differ. To standardize the level of detail across all feature descriptors, each column of the SUSC descriptor is treated as a discrete distribution. The distance between corresponding columns is then computed, with the virtual mapping shifted continuously to the right. Cosine similarity is employed for this calculation. The minimum distance is selected as the final result, and the yaw angle is determined by multiplying the number of shifts by the horizontal angular gap, as shown in Equation (6)

$$\begin{cases} dis = \min_{i=1}^{N_h^M} \frac{1}{N_h^L} \sum_{j=1}^{N_h^L} \left(\frac{F_{g(i,j)}^M \cdot F_j^L}{||F_{g(i,j)}^M|| \cdot ||F_j^L||} \right) \\ g(i,j) = (j+i)\% N_h^M \\ yaw = \frac{i \cdot 360}{N_h^M} \end{cases} \tag{6}$$

where N_h^L and N_h^M denote the number of columns of the point cloud's and virtual mapping's SUSC descriptor, respectively. dis denotes the final distance between the point cloud and virtual mapping. $g(i,j)$ denotes the index of column j after moving column i to the right, and yaw denotes the initial yaw angle.

After calculating the distance between the roadside LiDAR and the virtual mapping, a direct strategy is employed, based on a simple comparison of the distance differences derived from the minimum distance. Although straightforward, this method is susceptible to noise and local anomalies, potentially compromising its robustness. To enhance matching accuracy and stability, the similarity between virtual mappings is also considered, and the optimal virtual mapping result is computed based on the distance matrix.

Specifically, the geometric distance between each frame of the virtual mapping and the real LiDAR is first computed using multiple distance metrics. Subsequently, based on these distance values, all virtual mappings are ranked, and the top K candidate positions are selected, ensuring that those closest to the real LiDAR point cloud are retained. To further improve matching robustness, clustering analysis is applied. These candidate mappings are grouped into clusters based on their Euclidean distances, ensuring that mappings with high similarity are grouped together, thereby preventing instability in the matching results caused by a single anomalous candidate mapping. Each cluster represents a group of virtual mappings exhibiting similar geometric features.

Finally, the representativeness of each cluster is assessed by counting the number of elements within it. Within each cluster, the similarity between each element is computed to ensure consistency in the geometric features of the mappings. Additionally, to comprehensively evaluate the reliability of each cluster, the similarity between the real LiDAR point cloud and each element is considered. Using these similarity values, a comprehensive distance matrix is constructed, capturing the overall confidence level of each cluster, as shown in Equation (7)

$$DM = \begin{bmatrix} dis_{L1} & dis_{L2} & \cdots & dis_{LK} \\ dis_{11} & dis_{12} & \cdots & dis_{1K} \\ dis_{21} & dis_{22} & \cdots & dis_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ dis_{K1} & dis_{K2} & \cdots & dis_{KK} \end{bmatrix} \tag{7}$$

where dis_{Li} denotes the distance between the LiDAR point cloud and the i-th virtual mapping, and dis_{ij} denotes the distance between the i-th virtual mapping and the j-th virtual mapping. Based on this distance matrix, we calculate the distance between each row and the first row in the matrix, and take the virtual viewpoint position closest to the LiDAR point cloud as the final localization result.

3.4 Transformation Parameter Calculation

Once the correspondence between the Roadside LiDAR point cloud and virtual mapping is determined through feature retrieval, the transformation parameters can be estimated. First, we apply an initial transformation to the LiDAR point cloud based on the X, Y coordinates from the virtual mapping and the yaw angle obtained from feature retrieval. Next, we employ the Generalized Iterative Closest Point (GICP) algorithm to achieve precise registration between the LiDAR point cloud and the HD map, thereby estimating the extrinsic parameters.

4. Experiments

4.1 Experimental Setup

To comprehensively assess the performance of the proposed MapCalib framework, we conducted experiments using three distinct datasets: RLiDAR-sim, which simulates various scenarios; WHU-Urban3D Han et al. (2024), an outdoor dataset; and a Real-World dataset. The calibration effectiveness was quantitatively evaluated using two key metrics: Relative Rotation Error (RRE) and Relative Translation Error (RTE) Geiger et al. (2012), as expressed in Equation (8)

$$\begin{cases} \text{RRE} = \sum_{i=1}^{3} |F(R_T^{-1}, R_E)(i)| \\ \text{RTE} = ||t_T - t_E|| \end{cases} \tag{8}$$

where R_T and R_E denote the ground truth rotation matrix and the estimated rotation matrix, respectively. F denotes the function to calculate the Euler angle between two rotation matrices. RRE denotes the relative rotation error. Similarly, t_T and t_E denote the ground truth translation vector and the estimated translation vector, respectively. RTE denotes the relative translation error.

We compare the proposed method with several different LiDAR calibration algorithms: (1) ICP (Besl and McKay, 1992), which estimates the transformation matrix by iteratively minimising the distance between two point clouds; (2) GICP (Segal et al., 2009), which is based on a globally optimized version of ICP, and uses the covariance matrix to compute the objective function; (3) NDT (Biber and Straßer, 2003), which estimates the transformation matrix by transforming the point cloud data into a Gaussian distribution; (4) VI-eye (He et al., 2021), which achieves real-time registration of two point clouds by detecting a set of key semantic objects and basing it on their intrinsic shapes; and (5) OpenCalib (Wei et al., 2024), which uses a neural network to perform a rough calibration and then optimises it using an octree.

For the baseline algorithms, ICP, GICP, and NDT were implemented using the Point Cloud Library (PCL), and given that their performance is highly dependent on the initial position, they were tested using the initial position provided by GNSS signals. VIeye, in contrast, was implemented based on the referenced paper, with semantic segmentation performed using the RangeNet++ network for training. OpenCalib was evaluated using the provided open-source code. The parameters of these algorithms were carefully tuned to optimize performance on our dataset.

https://doi.org/10.5194/isprs-annals-X-G-2025-231-2025 | © Author(s) 2025. CC BY 4.0 License.

Method	*RLiDAR-sim*		*WHU-Urban3D*		*Real-World*	
	RRE°	$RTE(m)$	RRE°	$RTE(m)$	RRE°	$RTE(m)$
GNSS + ICP	5.28	0.48	2.98	1.03	15.88	2.36
GNSS + GICP	2.37	<u>0.20</u>	<u>2.07</u>	0.35	22.78	0.85
GNSS + NDT	13.14	0.84	7.44	1.83	13.16	4.12
VI-EYE	<u>2.54</u>	0.34	2.08	<u>0.33</u>	<u>6.79</u>	0.75
OpenCalib	10.54	0.62	8.34	0.63	8.62	<u>0.39</u>
Ours	**1.11** (+56.3%)	**0.127** (+36.8%)	**1.59** (+23.2%)	**0.277** (+15.3%)	**4.08** (+39.9%)	**0.241** (+39.3%)

Table 1. Quantitative results of calibration on three datasets. The best result is indicated in bold, while the second-best result is underlined.

4.2 Datasets

RLiDAR-sim: Due to the lack of publicly available datasets for roadside LiDAR calibration, RLiDAR-sim was developed using the ROS-GAZEBO simulation platform. This environment facilitates the deployment of multiple LiDAR sensors, the creation of various road and traffic scenarios, and the provision of accurate ground-truth data, making it ideal for testing algorithms. The simulation employed different LiDAR configurations to mimic real-world conditions, including two 128-line mechanical LiDAR with a 120-meter detection range: one with a 360-degree horizontal field of view (FOV) and another with a 120-degree horizontal FOV, both maintaining the same vertical FOV of -30 to 10 degrees to ensure comprehensive road coverage. The dataset was generated by initially scanning the entire scene with a vehicle-mounted 64-line mechanical LiDAR, which was then used to create an HD map. ROS was utilized to collect and synchronize the LiDAR data with the simulation environment, ensuring accurate sensor data for testing.

WHU-Urban3D: This is a large-scale, multi-source 3D point cloud dataset, covering over 3.2×10^6 square meters across Shanghai and Wuhan, was acquired using AS-900HL and HiScan-Z laser measurement systems. It includes both aerial laser scanning (ALS) and mobile laser scanning (MLS) point clouds, along with panoramic images and annotations for more than 200 million points. For this experiment, we focused on the MLS data, which were used as HD maps. A virtual reprojection model was manually created to simulate roadside LiDAR data generation, using a 128-line mechanical LiDAR sensor with a 200-meter detection range, a 120-degree horizontal FOV, and a vertical FOV from -30 to 10 degrees, reflecting common roadside sensor configurations. To enhance the realism of the simulation, Gaussian noise was added to replicate real-world LiDAR noise characteristics. The sensor's extrinsic parameters were manually adjusted, ensuring accurate alignment with the virtual environment for realistic data generation and subsequent testing.

Real-World Dataset: To evaluate the performance of MapCalib in real-world, a dataset was collected in a park in Wuhan, covering 7 intersections, including straight roads, "T" intersections, and crossroads. A total of 11 LiDAR sensors were deployed, capturing around 400 frames per sensor, with each frame containing approximately 60,000 points. The primary sensors used were (1) LS-LiDAR-C32, a mechanical LiDAR with a 150-meter detection range, a 360-degree horizontal FOV for full surrounding coverage, and a vertical FOV from -16 to 15 degrees, and (2) CH128X1, a semi-solid-state LiDAR with a 200-meter range, a 120-degree horizontal FOV focusing on the roadside and intersections, and a vertical FOV from -18 to 7 degrees. The HD map was created through mobile mapping vehicles equipped with LiDAR, inertial guides, RTK, and panoramic cameras, with data processed manually. Sensor positions were determined using a high-precision GNSS-RTK device.

4.3 Results and Analysis

We first quantitatively compared the calibration performance of the proposed MapCalib method with several baseline methods. Table 1 shows the results of the comparison of calibration errors. As shown, MapCalib achieved relatively accurate calibration, with a translation error under 0.3 m and a rotation error below 5 degrees. Notably, in the simulated environment and WHU-Urban3D dataset, most calibration methods performed well due to the relatively simple scenes and low noise levels. Although MapCalib achieved the best calibration results, its advantage was marginal. In real-world scenarios, however, all baseline methods struggled to achieve accurate calibration, while MapCalib demonstrated a notable advantage, successfully aligning data from different LiDAR systems. MapCalib showed a 25.2% improvement in RRE and a 39.3% improvement in RTE over the best baseline method.

It is evident that for traditional methods such as ICP, GICP, and NDT, the initial pose provided by GNSS (with an error typically around 5 meters) is insufficient for achieving high-precision calibration of roadside LiDAR. VI-eye enhances robustness by incorporating semantic information, thereby providing stronger constraints between different types of LiDAR. However, the performance of VI-eye is strongly dependent on the accuracy of the saliency point extractor, which is sensitive to low-density or high-noise point clouds, thereby limiting its broader applicability in real-world scenarios. OpenCalib, a calibration method designed for onboard LiDAR systems, experiences a notable decline in calibration accuracy when directly applied to roadside LiDAR.

Subsequently, we analyzed the relationship between computational efficiency and accuracy for the method proposed in this paper, as shown in the table 3. As the sampling interval increases, the localization error shows a progressively rising trend, whereas computation time exhibits an inverse relationship with the interval. Specifically, as the sampling interval increases, computation time decreases substantially. Notably, when the sampling interval reaches approximately 2 meters, a significant reduction in error is observed up to this point, with

https://doi.org/10.5194/isprs-annals-X-G-2025-231-2025

Dataset	*Rotation(deg)*			*Translation(m)*		
	Pitch	*Roll*	*Yaw*	*X*	*Y*	*Z*
RLiDAR-sim 1	0.89	2.10	2.87	0.04	0.17	0.03
RLiDAR-sim 2	1.24	2.57	3.56	0.10	0.18	0.10
WHU-Urban3D 1	2.11	3.12	5.13	0.20	0.29	0.14
WHU-Urban3D 2	1.86	4.21	4.54	0.29	0.26	0.14
Real-World 1	1.23	2.16	8.63	0.23	0.30	0.11
Real-World 2	1.42	2.64	7.95	0.12	0.38	0.10

Table 2. Calibration results for global position

further reductions leading to only marginal increases in processing time. When the sampling interval is reduced below 2 meters, the further reduction in error becomes negligible, while the increase in processing time becomes more pronounced.

Interval/(m)	Times/(s)	Error/(m)
0.5	2.53	0.38
1	1.46	0.53
2	0.61	0.77
3	0.54	1.8
4	0.60	4.85
5	0.52	9.72

Table 3. The impact of sampling interval on localization accuracy and computational efficiency

We further analyzed the relationship between sensors and the absolute geographic coordinate system. All baseline methods failed to achieve accurate global pose estimation due to the absence of an effective connection to the absolute geographic coordinate system. In contrast, the proposed MapCalib method substantially enhances the sensor's global georeferencing accuracy. By integrating high-precision mapping data, it enables reliable localization in complex urban settings and dynamic, rapidly changing environments. Table 2 shows calibration results for three datasets.

Finally, multiple LiDARs were calibrated in real environments, with results at an intersection displayed in Figure 3. The figure demonstrates that despite the low overlap and the inclusion of various LiDAR types, these factors nearly rendered baseline methods ineffective for valid calibration. In contrast, by incorporating HD map technology as constraints, MapCalib bridges data from different sensors. Additionally, the SUSC descriptors leverage sensor data features effectively, bridging differences across sensors and showcasing MapCalib's wide applicability and clear advantages in real-world applications.

5. Conclusion

This paper proposes a roadside LiDAR automatic calibration method aided by HD map, aimed at addressing the challenges of low overlap and viewpoint differences encountered. Specifically, to address the environmental representation discrepancies caused by viewpoint differences between roadside LiDAR

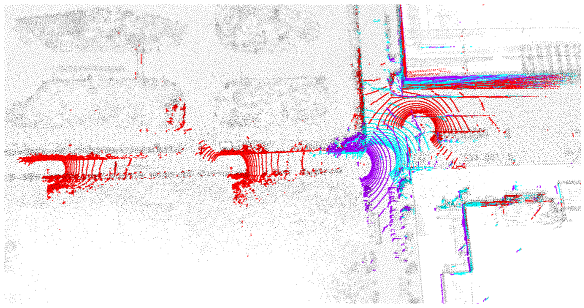

Figure 3. Calibration results of multiple LiDARs in a real scenario

and HD map, we first design an innovative virtual map projector that constructs a virtual mapping of the HD map from the roadside LiDAR perspective, thereby enhancing data similarity. Next, considering the installation modes and physical mechanisms of roadside LiDAR, we introduce a SUSC descriptor to ensure compatibility across different types of LiDAR. Finally, through feature retrieval and iterative optimization, we achieve high-precision calibration of roadside LiDAR. Extensive experiments across multiple datasets validate the effectiveness of our method, demonstrating that it can autonomously perform geographical calibration of roadside LiDARs even in the absence of prior information.

In the future, we will further explore real-time online calibration technologies for roadside LiDARs. Additionally, we plan to investigate the scalability of the method in dynamic sensing applications, promoting the overall advancement of environmental perception in vehicle-infrastructure cooperation and autonomous driving technologies.

References

Besl, P. J., McKay, N. D., 1992. Method for registration of 3-d shapes. *Sensor fusion IV: control paradigms and data structures*, 1611, Spie, 586–606.

Biber, P., Straßer, W., 2003. The normal distributions transform: A new approach to laser scan matching. *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*, 3, IEEE, 2743–2748.

Chiang, K.-W., Srinara, S., Tsai, S., Lin, C.-X., Tsai, M.-L., 2023. High-Definition-Map-Based LiDAR Localization Through Dynamic Time-Synchronized Normal Distribution Transform Scan Matching. *IEEE Transactions on Vehicular Technology*, 72(6), 7011–7023.

Geiger, A., Lenz, P., Urtasun, R., 2012. Are we ready for autonomous driving? the kitti vision benchmark suite. *2012 IEEE conference on computer vision and pattern recognition*, IEEE, 3354–3361.

Guo, Y., Zhou, J., Li, X., Tang, Y., Lv, Z., 2024. A Review of Crowdsourcing Update Methods for High-Definition Maps. *ISPRS International Journal of Geo-Information*, 13(3), 104.

Hakak, S., Gadekallu, T. R., Maddikunta, P. K. R., Ramu, S. P., M, P., De Alwis, C., Liyanage, M., 2023. Autonomous vehicles in 5G and beyond: A survey. *Vehicular Communications*, 39, 100551.

Han, X., Liu, C., Zhou, Y., Tan, K., Dong, Z., Yang, B., 2024. WHU-Urban3D: An urban scene LiDAR point cloud dataset for semantic instance segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 209, 500–513.

He, Y., Ma, L., Jiang, Z., Tang, Y., Xing, G., 2021. Vi-eye: semantic-based 3d point cloud registration for infrastructure-assisted autonomous driving. *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, 573–586.

Herau, Q., Piasco, N., Bennehar, M., Roldao, L., Tsishkou, D., Migniot, C., Vasseur, P., Demonceaux, C., 2024. Soac: Spatio-temporal overlap-aware multi-sensor calibration using neural radiance fields. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15131–15140.

Jiao, J., Liao, Q., Zhu, Y., Liu, T., Yu, Y., Fan, R., Wang, L., Liu, M., 2019. A novel dual-lidar calibration algorithm using planar surfaces. *2019 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 1499–1504.

Khan, M. U., Zaidi, S. A. A., Ishtiaq, A., Bukhari, S. U. R., Samer, S., Farman, A., 2021. A comparative survey of lidar-slam and lidar based sensor technologies. *2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC)*, IEEE, 1–8.

Kim, S.-W., Qin, B., Chong, Z. J., Shen, X., Liu, W., Ang, M. H., Frazzoli, E., Rus, D., 2014. Multivehicle cooperative driving using cooperative perception: Design and experimental validation. *IEEE Transactions on Intelligent Transportation Systems*, 16(2), 663–680.

Kulawiak, M., 2024. Comparison of 3D Point Cloud Completion Networks for High Altitude Lidar Scans of Buildings. *Photogrammetric Engineering & Remote Sensing*, 90(1), 55–64.

Lee, W., Won, C., Lim, J., 2020. Unified calibration for multi-camera multi-lidar systems using a single checkerboard. *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 9033–9039.

Li, L., Kong, X., Zhao, X., Huang, T., Li, W., Wen, F., Zhang, H., Liu, Y., 2021. Ssc: Semantic scan context for large-scale place recognition. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2092–2099.

Li, Y., Ibanez-Guzman, J., 2020. Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems. *IEEE Signal Processing Magazine*, 37(4), 50–61.

Liao, Y., Li, J., Kang, S., Li, Q., Zhu, G., Yuan, S., Dong, Z., Yang, B., 2023. SE-Calib: Semantic Edge-Based LiDAR–Camera Boresight Online Calibration in Urban Scenes. *IEEE Transactions on Geoscience and Remote Sensing*, 61, 1–13.

Liu, R., Wang, J., Zhang, B., 2020. High definition map for automated driving: Overview and analysis. *The Journal of Navigation*, 73(2), 324–341.

Ren, H., Zhang, S., Li, S., Li, Y., Li, X., Ji, J., Zhang, Y., Zhang, Y., 2023. Trajmatch: Toward automatic spatio-temporal calibration for roadside lidars through trajectory matching. *IEEE Transactions on Intelligent Transportation Systems*.

Segal, A., Haehnel, D., Thrun, S., 2009. Generalized-icp. *Robotics: science and systems*, 2number 4, Seattle, WA, 435.

Srinara, S., Chiu, Y.-T., Tsai, M.-L., Chiang, K.-W., 2022. High-definition point cloud map-based 3D LiDAR-IMU calibration for self-driving applications. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 271–277.

Wei, P., Yan, G., Li, Y., Fang, K., Cai, X., Liu, W., Yang, J., 2024. Automatic multi-LiDAR calibration and refinement method. *Measurement Science and Technology*, 35(7), 075203.

Wong, H., Zhang, X., Wen, T., Yang, M., Jiang, K., Yang, D., 2021. Hd-map aided lidar-ins extrinsic calibration. *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, IEEE, 3136–3143.

Xiao, J., Wang, S., Zhou, J., Tian, Z., Zhang, H., Wang, Y.-F., 2024. MIM: High-Definition Maps Incorporated Multi-View 3D Object Detection. *IEEE Transactions on Intelligent Transportation Systems*, 1-13.

Xie, Y., Shao, R., Guli, P., Li, B., Wang, L., 2018. Infrastructure based calibration of a multi-camera and multi-lidar system using apriltags. *2018 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 605–610.

Zhang, C., Wei, J., Hu, A. S., Fu, P., 2022. A novel method for calibration and verification of roadside millimetre-wave radar. *IET Intelligent Transport Systems*, 16(3), 408–419.

Zhang, G., Wu, K., Lin, J., Wang, T., Liu, Y., 2024. Automatic Extrinsic Parameter Calibration for Camera-LiDAR Fusion using Spherical Target. *IEEE Robotics and Automation Letters*.

Zhang, Z., Zheng, J., Sun, R., Zhang, Z., 2020. 3d point cloud registration for multiple roadside lidars with retroreflective reference. *2020 IEEE International Conference on Networking, Sensing and Control (ICNSC)*, IEEE, 1–6.

Zhao, C., Shi, Y., Du, Y., Jiang, S., Ji, Y., Zhao, X., 2024. A rapid and convenient spatiotemporal calibration method of roadside sensors using floating connected and automated vehicle data. *IEEE Transactions on Intelligent Transportation Systems*.

Zhou, L., Li, Z., Kaess, M., 2018. Automatic extrinsic calibration of a camera and a 3d lidar using line and plane correspondences. *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 5562–5569.