The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIX-B1-2026
XXV ISPRS Congress 2026 "From Imagery to Understanding", Commission I, 4–11 July 2026, Toronto, Canada

Loose Coupling Modeling of LiDAR-based Localization and SLAM

Chenglu Wen[1,2], Huanjia Zhang[1,2], Cheng Wang[1,2]

[1] Fujian Key Laboratory of Urban Intelligent Sensing and Computing, Xiamen University, China
[2] Key Laboratory of Multimedia Trusted Perception and Efficient Computing, Xiamen University, China

Keywords: LiDAR-based localization, scene coordinate regression, simultaneous localization and mapping, loose coupling

Abstract

In recent years, LiDAR-based localization has been widely explored. Among them, Scene Coordinate Regression (SCR)-based methods have demonstrated outstanding accuracy and robustness in city scenes. Integrating these models with traditional Simultaneous Localization and Mapping (SLAM) methods is expected to enhance localization accuracy and reliability further. This paper proposes loosely coupled fusion methods integrating an SCR model with SLAM to improve localization accuracy and robustness. The approach addresses the information loss problem in high-level sensor fusion while maintaining computational efficiency. The method achieves tighter data association and complementary performance advantages by strategically combining LiDAR-based localization results with SLAM pose estimates. Experimental results in the NCLT and HeLiPR datasets demonstrate that the proposed fusion framework effectively corrects SLAM drift and maintains stable pose estimation accuracy under diverse environmental conditions. Furthermore, the sparse-frame coupling strategy significantly reduces computational overhead without degrading localization performance, making the method suitable for practical applications. The system exhibits improved robustness across regions and LiDAR configurations while preserving real-time operation capabilities.

1. Introduction

Localization technology serves as a fundamental enabler for autonomous systems, playing a critical role in the safe navigation and operation of self-driving vehicles. Currently, mainstream intelligent passenger vehicles adopt a multi-sensor fusion localization framework, integrating the Global Navigation Satellite System (GNSS), Inertial Navigation System (INS), and visual perception to support localization in complex scenarios in a complementary manner. This approach has proven effective in most conventional urban environments. However, widespread GNSS-denied conditions—resulting from physical obstructions or electromagnetic interference—pose significant challenges to existing localization systems.

During GNSS outages, inertial navigation and vision-based large-scale feature matching often introduce substantial errors. In contrast, LiDAR captures precise 3D spatial data, providing advantages such as all-weather operation and high geometric measurement accuracy. This makes it a crucial solution for navigation in GNSS-denied environments.

At present, LiDAR-based visual localization for autonomous vehicles primarily employs two approaches: (1) Relative localization via LiDAR-inertial odometry, though this method suffers from cumulative errors and lacks global referencing lines. (2) Absolute localization employs neural networks that regress global 6-DoF poses from 3D point cloud sequences, despite the associated computational overhead.

To address two critical challenges in GNSS-denied environments: (1) the limited feature representation capacity and computational inefficiency of SCR methods, and (2) the error accumulation in long-range SLAM localization, this paper proposes a loosely-coupled framework that integrates LiDAR-visual localization models with conventional SLAM. As shown in Figure 1, The LiDAR-visual component employs the extensively validated SGLoc (Li et al., 2023a) and its efficient derivative,

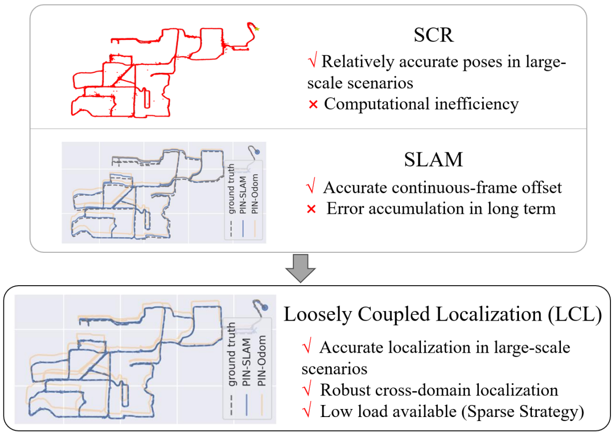


Figure 1. Illustration of the strengths and weaknesses between SCR and SLAM. It shows the advantages and disadvantages of SCR and SLAM. This paper proposes a loosely coupled localization method (LCL) to integrate both advantages and compensate for their shortcomings.

LightLoc (Li et al., 2025) for rapid pretraining, while the SLAM backbone utilizes the state-of-the-art PIN-SLAM. The proposed multi-mode loose-coupling paradigm is evaluated on large-scale urban datasets NCLT and HeLiPR. The main contributions of this paper are as follows:

- A loosely coupled fusion framework of LiDAR vision, SCR model, and SLAM is designed to complement the advantages of the two technologies.
- We propose a sparse frame loose coupling strategy to reduce computational overhead while ensuring positioning accuracy.
- For different LiDAR configurations and environmental scenarios, the cross-domain adaptation ability of the framework is verified through experiments.

2. Related Work

In GNSS-denied environments, LiDAR-based visual perception and localization have become essential solutions for the navigation of autonomous vehicles. This technology can be categorized into two primary approaches: explicit map-based and implicit map-based methods.

2.1 Explicit Map-Based Localization

The explicit map-based localization methods achieve positioning by estimating the relative transformation between input data and pre-built maps. For instance, DiSCO (Xu et al., 2021) utilizes point cloud projection and multi-layer bird's-eye view (BEV) representations to estimate location. BEVPlace (Luo et al., 2023) employs group convolution and NetVLAD (Arandjelović et al., 2018) for feature extraction and place recognition, leveraging a lightweight convolutional neural network to process the BEV image representation of LiDAR data. It achieves accurate global localization through place recognition, followed by a three-degree-of-freedom (3-DoF) pose estimation. BEVPlace++ (Luo et al., 2025), an improved version, incorporates a rotation-equivariant module (REM) to obtain distinctive features while enhancing robustness against rotational variations.

2.2 Implicit Map-Based Localization

The implicit map-based localization approaches infer poses directly from input data via neural networks, utilizing methods such as direct pose regression and scene coordinate regression (SCR). PointLoc (Wang et al., 2022) employs PointNet++ (Qi et al., 2017) with self-attention mechanisms for end-to-end pose regression. SGLoc efficiently encodes geometric priors for SCR, followed by RANSAC-based pose estimation. LiSA (Li et al., 2023b) enhances SCR by incorporating semantic awareness through the distillation of knowledge from segmentation models. LightLoc (Li et al., 2025) accelerates the learning process by freezing scene-agnostic feature backbones while fine-tuning scene-specific prediction heads.

2.3 Simultaneous Localization and Mapping

SLAM (Simultaneous Localization and Mapping) is a technology that enables robots and other unmanned systems to construct maps in real time while determining their positions within unknown environments through sensor measurements. It is a fundamental technology in fields such as autonomous driving, UAV navigation, industrial robotics, and intelligent devices. Currently, the predominant SLAM technologies are LiDAR-based SLAM (utilizes LiDAR sensors), visual SLAM (employs monocular or binocular cameras) and multi-sensor fusion SLAM. Regarding implementation, SLAM is mainly divided into Filter-Based SLAM and Graph-Based SLAM. PIN-SLAM (Pan et al., 2024) is a SLAM method based on point-based implicit neural map representation, which achieves globally consistent loop closure detection and mapping. Its front-end is a LiDAR-based Visual Odometry (VO), and the back-end adopts graph optimization. The optimized neural point local map is used for global loop closure detection and correction, and Pose Graph Optimization (PGO) is performed to correct cumulative errors.

3. Method

By loosely coupling the LiDAR visual positioning model SCR with the SLAM model, we effectively integrate the complementary strengths of both technologies to minimize positioning errors and enhance system robustness. The proposed method consists of two components: dense loose coupling and sparse loose coupling. Through dense loose coupling, we investigate whether introducing the LiDAR visual positioning model can mitigate the accumulation of SLAM's offset error and reduce the occurrence of outlier pose estimations. Sparse loose coupling, on the other hand, aims to explore the feasibility of decreasing the positioning frequency of the LiDAR visual positioning model to reduce computational load while maintaining acceptable positioning accuracy under the limited computing resources typically available in real-world vehicles.

The visual localization system based on scene coordinate regression (SCR) primarily employs the SGLoc model. The SGLoc method achieves high-precision localization by efficiently encoding scene geometric information. Building upon this foundation, LiSA adopts a teacher-student model that transfers knowledge from the segmentation model to the coordinate regression network through knowledge distillation, allowing the model to focus on categories that are beneficial for localization. LightLoc accelerates the learning process by freezing the scene independent feature backbone network and training only scene-specific prediction heads. Moreover, Sample Classification Guided (SCG) and Redundant Sample Distillation (RSD) techniques have been introduced to reduce ambiguity caused by visually similar regions and to alleviate the computational burden during the training process, thereby enhancing overall training efficiency.

The SLAM method implemented in our model is based on PIN-SLAM (Pan et al., 2024). PIN-SLAM utilizes a point-based implicit neural map representation, enabling globally consistent loop closure detection and mapping. The front-end employs LiDAR-based visual odometry (VO), while the back-end incorporates graph optimization. An optimized local map of nerve points facilitates global loop closure detection and correction, and pose graph optimization (PGO) is conducted to address cumulative errors.

The loosely coupled strategy is applied during the backend pose optimization, as illustrated in Figure 2. Once the visual localization model retrieves the predicted world coordinates for each point, the pose transformation of each frame point cloud relative to the initial point is computed using the Random Sample Consensus algorithm (RANSAC) for optimization. According to different screening strategies, the pose predictions generated by the visual localization model are dynamically integrated into the back-end factor graph. Simultaneously, based on the prediction results, the covariance matrix of the pose prior factor is dynamically adjusted, thereby enhancing the system's robustness and accuracy.

3.1 Loosely Coupled Localization For Dense Frames

In the dense loosely coupled method, we perform a coarse filtering by comparing the differences between consecutive frames of the pose predicted by the visual localization model and the SLAM prior pose. Specifically, a sliding window of size l is used to calculate the pose differences between the current frame's SCR pose and the preceding and subsequent frames within the window, which are summed up as $diff_{SCR}$. Similarly, the pose differences between the current frame's SLAM prior pose and the preceding and subsequent frames within the window are calculated and summed up as $diff_{SLAM}$. If the translational and rotational errors of $diff_{SCR}$ and $diff_{SLAM}$ fall within the defined threshold ranges (1.0 m for translation and 0.1 rad for

https://doi.org/10.5194/isprs-archives-XLIX-B1-2026-201-2026

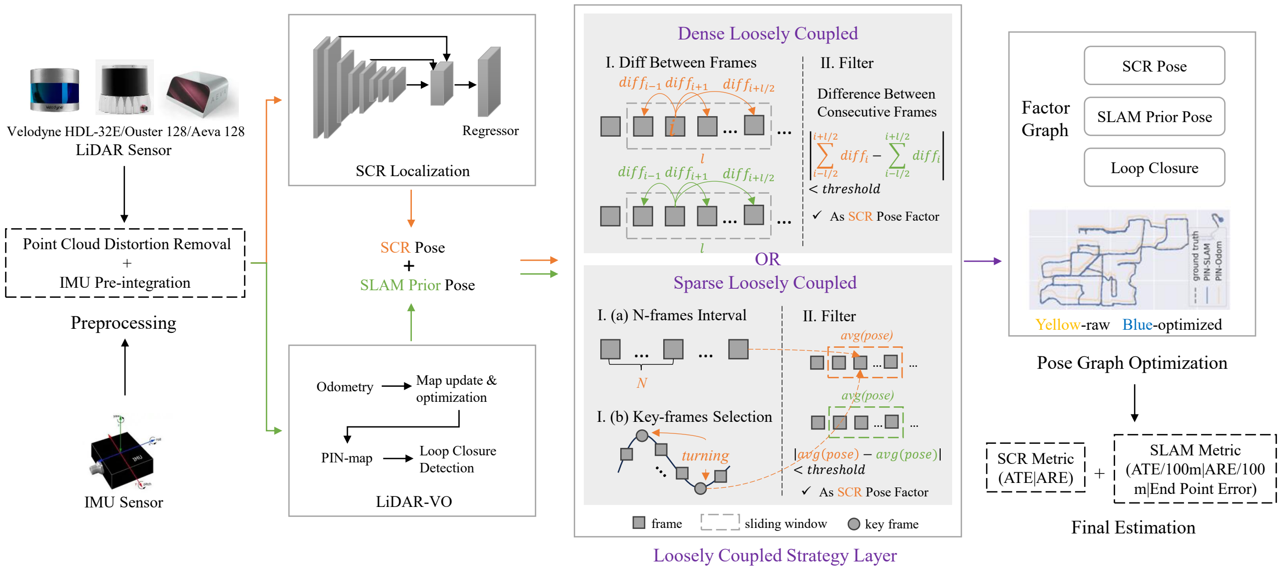


Figure 2. The pipeline of the proposed loosely coupled localization method (**LCL**).

rotation), the SCR pose estimation is considered to have high confidence and is selected for inclusion in the backend factor graph.

3.2 Loosely Coupled Localization For Sparse Frames

Beginning with the current state of autonomous driving localization systems, frequent and high-rate queries to the visual localization model for positioning can considerably elevate the processing load on the onboard chip. Consequently, it is crucial to develop various sparse localization strategies to minimize the frequency of calls to the visual localization model for positioning. In particular, two strategies can be identified: (a) N-frames Interval, and (b) Key-frames Selection. Subsequently, further filtering is performed afterward.

N-frames Interval and Filtering Strategy A key pose is extracted every N frame.

Key(Turning)-frames Selection and Filtering Strategy Although turning movements occur less frequently in the trajectory (less than 10%), they can cause significant pose changes due to environmental interference. Therefore, strategically incorporating turning key points into the factor graph enables substantial improvement in localization accuracy with minimal computational overhead. Given that turns develop gradually (demonstrating small inter-frame rotation rates), we identify sustained rotational changes through:

Let the size of the sliding window be N. The rotation variation of the frame j within the window ($j = i - \lfloor \frac{N-1}{2} \rfloor, i - \lfloor \frac{N-1}{2} \rfloor + 1, ..., i + \lfloor \frac{N-1}{2} \rfloor$) is denoted as ΔR_j. Then, the formula for calculating the cumulative rotation variation $\sum |\Delta R_j|$ within the window is as follows:

$$\sum_{j=i-\lfloor \frac{N-1}{2} \rfloor}^{i+\lfloor \frac{N-1}{2} \rfloor} |\Delta R_j| \tag{1}$$

whereΔR_j represents the rotation change of the frame j relative to the $j-1$ frame, which can be calculated from the attitude angles of the two frames.

When $\sum_{j=i-\lfloor \frac{N-1}{2} \rfloor}^{i+\lfloor \frac{N-1}{2} \rfloor} |\Delta R_j| > T$ (where T represents the threshold of cumulative rotation change), a steering process is detected, and the central frame of the window is designated as the key steering frame. Then, we skip M frames to avoid over-marking.

3.3 Further Comparison And Filtering

After initially filtering out the sparse SCR localization poses, it is necessary to incorporate the SLAM prior poses for further comparison and filtering. Specifically, a sliding window of fixed size is used, with the current frame as the central frame. The difference between the average pose of the SCR localization model and the average pose of the SLAM prior within the window is compared. If the translational and rotational errors fall within the defined threshold ranges (1.0 m for translation and 0.1 rad for rotation), the pose estimation is considered to have high confidence and is selected to be added to the backend factor graph.

4. Experiments

Dataset Benchmarks. We trained and evaluated a LiDAR-based visual localization model using the loosely coupled SLAM approach on two large-scale localization datasets: NCLT (Nicholas et al., 2015) and HeLiPR (Jung et al., 2023). The datasets contain various weather, traffic, and lighting conditions. Ground truth pose is generated by the interpolations of INS+RTK. For the NCLT dataset, the sparse 32-line point cloud data acquired by the Velodyne HDL-32E sensor are selected. The data collected on 2012-01-22 (26,145 frames), 2012-02-02 (29,527 frames), 2012-02-18 (26,930 frames) and 2012-05-11 (25,496 frames) is used as the training set, totaling 108,098 frames, while the data from 2012-03-31 (26,326 frames) is designated as the test set.

The HeLiPR dataset incorporates dense 128-line point cloud data collected by two different sensors: the omnidirectional rotating Ouster 128 sensor and the narrow-field solid-state Aeva 128 sensor. The Ouster sensor provides a 360° panoramic field of view, whereas the Aeva sensor provides a forward-facing field of view of approximately 120°.

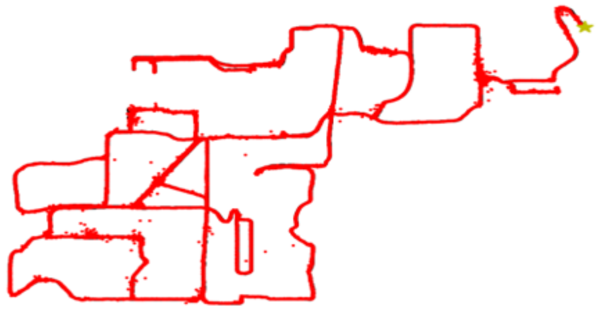


Figure 3. The visualization of SGLoc on NCLT.

Figure 4. The visualization of loosely coupled-localization on NCLT.

4.1 Loosely Coupled Localization For Dense Frames

Table 1 presents the localization results for SGLoc, PIN-SLAM, and the loosely cou-pled SGLoc and PIN-SLAM methods on the NCLT dataset. The evaluation metrics are divided into two categories: SCR and SLAM. The SCR metrics consist of the average translation error (ATE) and the average rotation error (ARE). The SLAM metrics include the average translation error per 100 meters (ATE/100m), the average rotation error per 100 meters (ARE/100m), and the translation error at the end-point.

As evidenced in Table 1, the cumulative error of the SLAM system has been rectified with notable precision through the poses provided by the loosely coupled visual localization model. Meanwhile, the outliers exhibiting large error deviations predicted by the visual localization system have been effectively corrected by the more precise continuous frame pose offset estimation offered by the SLAM system. Although the average translation error has increased marginally (by 0.25 meters), the average rotation error has seen a decrease (by 0.7 degrees). Significantly, the cumulative translation error at the end point of the SLAM system has reduced significantly (by 13.22 meters). Furthermore, the overall robustness of the loosely coupled system has been improved, eliminating any significant trajectory deviations.

Figure 3 shows the visualization of SGLoc on the NCLT dataset. And figure 4 shows the visualization of the localization results of the loosely coupled SGLoc and PIN-SLAM on the NCLT dataset. The yellow trajectory represents the registration-only output, while the blue trajectory illustrates the factor graph optimized result.

Methods	SCR Metrics		SLAM Metrics		
	ATE	ARE	ATE/100m	ARE/100m	End Point Error
SGLoc	1.706m	4.018°	\	\	\
PIN-SLAM	\	\	1.831cm	1.183°	16.129m
Loosely Coupled	1.976m	3.346°	2.233cm	1.407°	2.900m

Table 1. Localization results on the NCLT dataset.

4.2 Loosely Coupled Localization For Sparse Frames

Frame Skipping Selection Strategy The method skips frames at fixed intervals and attempts to incorporate loosely coupled localization results into the factor graph. If the pose predicted by the visual localization model differs from the SLAM prior pose by less than a specified threshold, it is incorporated into the backend factor graph.

Table 2 presents the loosely coupled localization results obtained using the frame skipping selection strategy. Figures 5(a), 5(b) and 5(c) illustrate the visual results. After incorporating sparse SCR poses (e.g., at 100-frame intervals) into the factor graph for joint optimization, the cumulative error of the SLAM system is significantly reduced, and any potential factors with large deviations are corrected. As illustrated in the figure, as the frame interval increases, the deviation between the localization trajectory and the ground truth trajectory increases slightly; however, the overall consistency remains strong. This outcome verifies the effectiveness of the sparse frame strategy in reducing computational overhead while ensuring localization accuracy. Even when the introduced SCR poses are extremely sparse (e.g., at 1000-frame intervals), a portion of the SLAM system's cumulative error is still eliminated.

Sampling Interval N	Frames Used M	ATE	ARE
100	262	2.496m	3.011°
500	52	4.321m	3.224°
1000	26	8.327m	2.893°

Table 2. The loosely coupled localization results obtained using the frame skipping Selection strategy.

Keyframe Selection Threshold Configuration		
keyframe_win_len	1.0	0.75
angle_thresh_rad	30	25
post_turn_skip_frames	100	100

Table 3. The configuration of key frame selection threshold.

Key Frame Selection Strategy This strategy's threshold settings include: (1) the observation window size for key frame identification, (2) a radian-based angular threshold, and (3) a frame skip mechanism triggered after turn detection to prevent redundant key frames in identical curved segments. Table 3 presents the configurations of key frame selection. Table 4 presents the results of key frame selection. The initial screening process selects key frames based on the threshold established at the top of the table. Subsequently, the refinement process involves comparing pose discrepancies between visual-SLAM predictions and prior poses across consecutive frames, thus eliminating erroneous outlier estimations. Table 5 shows the loosely-coupled localization results using the key frame screening strategy. By incorporating only the sparse SCR key frame poses into the factor graph for joint optimization, we significantly reduce the computational burden to a mere 0.3% of the original. Remarkably, even when the introduced SCR poses are extremely sparse (fewer than 100 frames), a portion of the cumulative error within the SLAM system is still mitigated. Figures 7(a) and 7(b) illustrate the loosely coupled localization results obtained using the key frame selection strategy.

The visualization of selected key frames in Figures 6(a) and 6(b) (marked with red dots) confirms that the proposed screening method effectively identifies turning-critical frames.

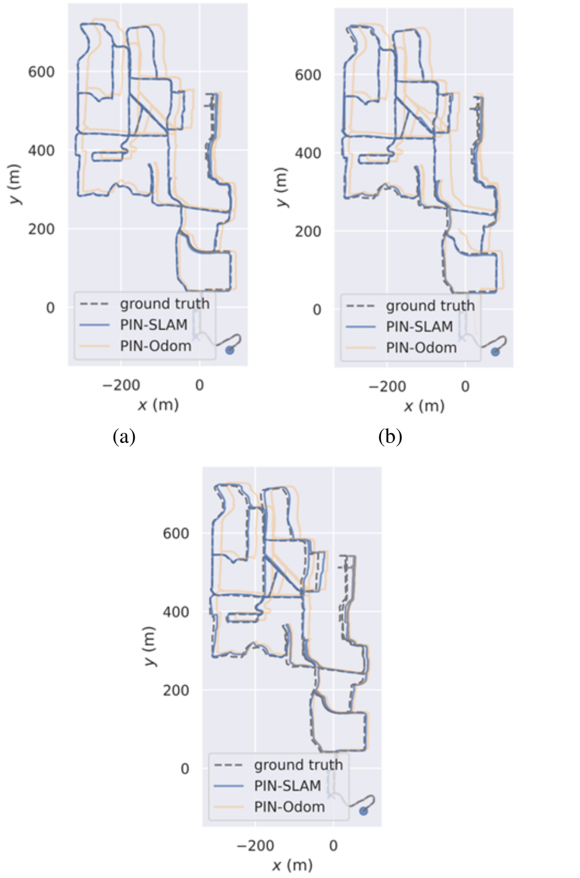


Figure 5. a, b, and c represent the visualizations of loosely coupled localization results at intervals of 100 frames (262 frames used), 500 frames (52 frames used), and 1000 frames (26 frames used).

Results of Key Frame Selection		
count after initial screening	80	98
count after refinement	64	93

Table 4. The count results of Key frame Selection.

The graphical results confirm that our key point screening-based loosely coupled approach significantly mitigates error accumulation in SLAM systems with sparse visual constraints (less than 100 frames). This approach fulfills the core requirements of reducing computational overhead while maintaining acceptable accuracy.

4.3 Cross-Domain Evaluation of Solid-State LiDAR Localization Transferability.

Our experiments initially utilized 32-beam sparse point clouds with complete omnidirectional coverage. To establish generalizability, we systematically evaluated the framework across: (1) varying LiDAR beam configurations (ranging from 32-beam to 128-beam) and (2) different viewing angle constraints. The adaptation process leveraged LightLoc’s two-stage paradigm by freezing the pre-trained backbone while fine-tuning the domain-specific prediction headers. Cross-domain validation was conducted on the HeLiPR benchmark, specifically focusing on the challenging DCC04-DCC06 sequences that present urban canyon effects.

Methods	ATE	ARE
64 frames after refinement	5.578m	3.391°
93 frames after refinement	3.588m	3.343°

Table 5. The loosely coupled localization results obtained using the key frame Selection strategy.

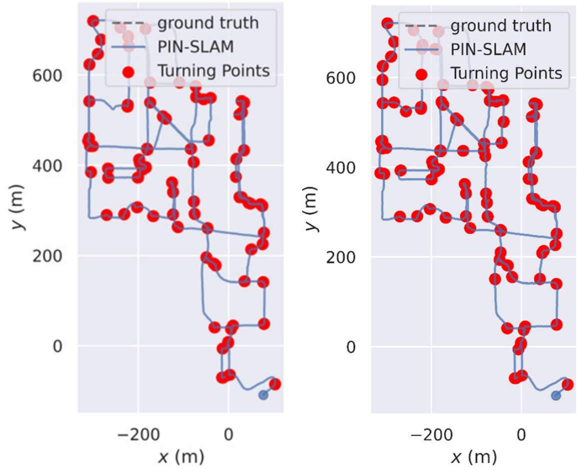


Figure 6. The visualization of selected key frames which are marked with red dots on NCLT.

We establish a cross-sensor evaluation protocol. The training set comprises 18,598 frames from the Ouster-128 (360° coverage) acquired in DCC04 and DCC06, while the test set includes 10,810 frames from the Aeva (120° frontal field of view) obtained in DCC05.

Unlike the NCLT dataset’s unidirectional trajectory design (all 4 trajectories), HeLiPR’s multidirectional paths create inherent orientation ambiguities. The cross-sensor domain shift exacerbates this issue, resulting in numerous trajectory-deviating outliers in the visual localization outcomes. Figure 8(a) illustrates the localization results of LightLoc on the DCC05 trajectory, while Figure 8(b) presents the localization results of PIN-SLAM on the same trajectory.

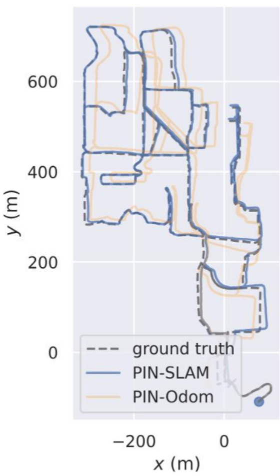
To address the degradation in cross-domain localization performance, we utilize a sliding window-based pose consistency validation mechanism. This approach selectively incorporates only high-confidence pose estimates into the pose graph optimization.

Methods	SCR Metric		SLAM Metric		
	ATE	ARE	ATE/100m	ARE/100m	End Point Error
LightLoc	12.78m	10.12°	\	\	\
PIN-SLAM	\	\	3.22cm	1.48°	43.23m
Loosely Coupled	3.72m	3.15°	3.91cm	1.53°	4.02m

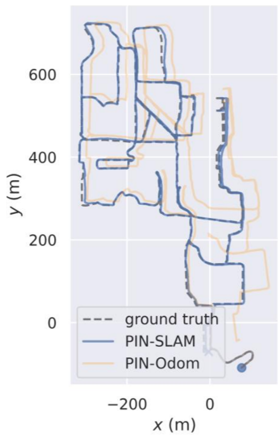
Table 6. Cross-domain localization results on HeliPR dataset.

As can be seen in Table 6, in the more challenging cross-domain small-view scenario, both the pure visual localization scheme and the pure SLAM scheme suffer from accuracy degradation. The average localization error of LightLoc exceeds 10 meters, while the cumulative error of PIN-SLAM results in a significant endpoint offset, with a translation distance exceeding 40 meters. In contrast, the loose-coupling scheme effectively filters out the erroneous poses provided by the visual localization model, utilizing the relative pose offsets from SLAM. Consequently, the

https://doi.org/10.5194/isprs-archives-XLIX-B1-2026-201-2026



(a)



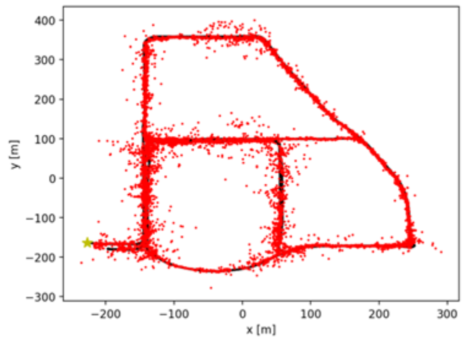
(b)

Figure 7. The visualization of selected key frames which are marked with red dots on NCLT.

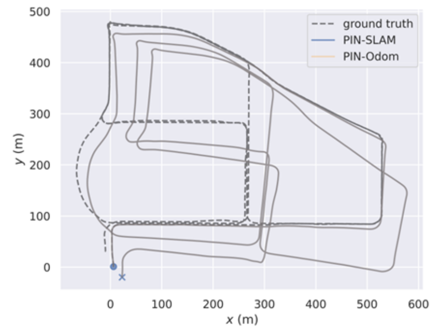
cumulative error of SLAM is rectified after incorporating the relatively accurate poses from the visual localization model into the factor graph. This joint optimization leads to a more accurate localization result. Specifically, the average translation error decreased from 12.78 meters to 3.72 meters, while the average rotation error reduced from 10.12 degrees to 3.15 degrees. Moreover, the average error per 100 meters remained essentially unchanged, with the final error decreasing from 43.23 meters to 4.02 meters. Figure 8(c) illustrates the localization results of loosely coupled localization on the DCC05 trajectory. In the figure, the yellow trajectory represents the registration-only approach, while the blue trajectory indicates the factor graph optimized outcome.

5. Conclusion

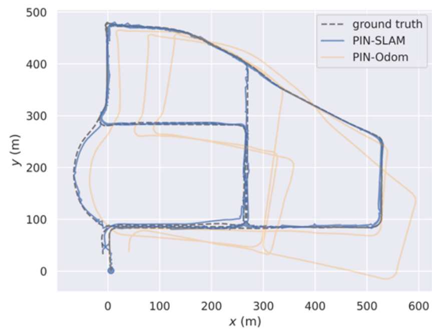
To address the low computational efficiency of LiDAR visual localization methods and the error accumulation issues encountered in traditional LiDAR and inertial fusion positioning, we propose a loose coupling method integrating SCR and SLAM positioning models (LCL). This method is specifically designed for scenarios characterized by sparse frames, and we have verified its feasibility for real-time operation within such a loosely coupled framework. Furthermore, cross-domain validation is conducted using datasets that encompass varying numbers of LiDAR lines and different viewpoints.



(a)



(b)



(c)

Figure 8. The visualization of LightLoc (a), PIN-SLAM (b) and loosely coupled-localization (c) on DCC05.

References

Arandjelović, R., Gronat, P., Torii, A., Pajdla, T., Sivic, J., 2018. NetVLAD: CNN Architecture for Weakly Supervised Place Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(6), 1437-1451.

Jung, M., Yang, W., Lee, D., Gil, H., Kim, G., Kim, A., 2023. Helipr: Heterogeneous lidar dataset for inter-lidar place recognition under spatial and temporal variations.

Li, W., Liu, C., Yu, S., Liu, D., Zhou, Y., Shen, S., Wen, C., Wang, C., 2025. Lightloc: Learning outdoor lidar localization at light speed. *Proceedings of the Computer Vision and Pattern Recognition Conference*, 6680–6689.

Li, W., Yu, S., Wang, C., Hu, G., Shen, S., Wen, C., 2023a. Sgloc: Scene geometry encoding for outdoor lidar localization. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9286–9295.

Li, W., Yu, S., Wang, C., Hu, G., Shen, S., Wen, C., 2023b. Sgloc: Scene geometry encoding for outdoor lidar localization. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9286–9295.

Luo, L., Cao, S.-Y., Li, X., Xu, J., Ai, R., Yu, Z., Chen, X., 2025. BEVPlace++: Fast, Robust, and Lightweight LiDAR Global Localization for Autonomous Ground Vehicles. *IEEE Transactions on Robotics*, 41, 4479-4498.

Luo, L., Zheng, S., Li, Y., Fan, Y., Yu, B., Cao, S.-Y., Li, J., Shen, H.-L., 2023. Bevplace: Learning lidar-based place recognition using bird’s eye view images. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 8666–8675.

Nicholas, C.-B., Arash, K. U., Ryan, M. E., 2015. University of Michigan North Campus long-term vision and lidar dataset. *IJRR*, 35, 545-565.

Pan, Y., Zhong, X., Wiesmann, L., Posewsky, T., Behley, J., Stachniss, C., 2024. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. *IEEE Transactions on Robotics*, 40, 4045-4064.

Qi, C. R., Yi, L., Su, H., Guibas, L. J., 2017. Pointnet++: deep hierarchical feature learning on point sets in a metric space. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, Curran Associates Inc., 5105–5114.

Wang, W., Wang, B., Zhao, P., Chen, C., Clark, R., Yang, B., Markham, A., Trigoni, N., 2022. PointLoc: Deep Pose Regressor for LiDAR Point Cloud Localization. *IEEE Sensors*, 22, 959-968.

Xu, X., Yin, H., Chen, Z., Li, Y., Wang, Y., Xiong, R., 2021. DiSCO: Differentiable Scan Context With Orientation. *IEEE Robotics and Automation Letters*, 6(2), 2791-2798.