2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Active Vehicle Re-localization Based on Non-repetitive LiDAR with Gimbal Motion Strategy

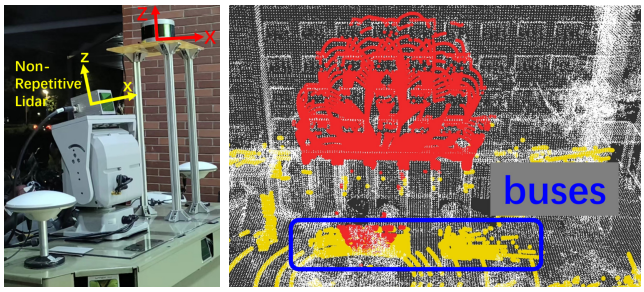
Xin'ao Wu, Chenxi Yang, Yiyang Guo, Hanyang Zhuang, Chunxiang Wang, Ming Yang

***Abstract*— The installation of a multi-layer 3D LiDAR atop the vehicle is a widely adopted hardware configuration for map-matching-based localization in intelligent driving. By offering a comprehensive 360° horizontal Field of View (FoV), this setup aims to achieve precise matching outcomes through the imposition of substantial geometric constraints against dynamic interferences and structural degradation. However, several factors limit its environmental adaptability, such as sparse point cloud density at distances, insufficient maximum sensing range, and notably, the restricted beam elevation angle, limiting the perception of the environment beyond obstacles. The rapid advancement of non-repetitive scanning LiDARs shows promise in mitigating such limitations. Nevertheless, their narrow FoV remains a challenge to overcome. In this study, we propose a solution by mounting such one single LiDAR on a two-axis rotating gimbal, enabling the vehicle to surpass the ranges and vertical FoV limitations of traditional setups actively. The corresponding gimbal motion strategy has been designed to automatically focus on the environment component with the most robust geometric constraints. Experimental results validate that the proposed method achieves superior robustness under high dynamic interference while delivering sufficient performance under standard conditions.**

I. Introduction

Map-matching-based vehicle localization in a pre-explored region, also known as re-localization, is widely utilized in contexts where Global Navigation Satellite System (GNSS) signals are unavailable. Ever since the early exploration of intelligent driving [1], LiDARs, designed primarily for environmental perception, have been commonly mounted on the vehicle roof to improve visibility. Re-localization, which also requires rich real-time environmental sensibility to match the pre-defined map, naturally shares such hardware configuration. As LiDAR technology has progressed, a single multi-layer 3D LiDAR mounted atop a vehicle roof has become the prevalent image associated with intelligent vehicles in the public consciousness.

However, re-localization and environmental perception ultimately constitute two distinct tasks, which is notably reflected in their management of dynamic obstacles. In environmental perception, detecting the vehicle's surroundings is sufficient to complete the task. However in terms of re-localization, dynamic obstacles not only act as disturbances but also obscure the background static environmental structures that should originally be perceived. Therefore, even if dynamic interferences are perfectly detected and filtered out, they may still cause serious repercussions to persist.



(a) (b)

Fig. 1. The proposed LiDAR configuration. (a) A non-repetitive LiDAR mounting on a two-axis rotating gimbal enlarges the FoV to the upper hemisphere conventionally as the blind spot while ensuring the structure density for registration, especially at a distance. (b) The scan comparison with the conventional setting in severe dynamic occlusion. Yellow: conventional 3D multi-layer LiDAR. Red: the proposed system.

To address this issue, two approaches become evident: enhancing the convergence ability of the remaining static environment structure or seeking additional structure out of the conventional Field of View (FoV). In our previous study [2] concerning the second approach, we positioned a hemispherical FoV LiDAR upward to mitigate the re-localization challenges caused by the absence of static features in the production lines of factory workshops. This deployment strategy was chosen to leverage the roof truss structure of the facility for matching.

Since such LiDARs were originally designed to offer blind spot coverage in front of vehicles, their performances have specialized in offering large FoV rather than obtaining effective sensing range or angular resolution. Consequently, that customization encountered limitations when generalized to other highly dynamic scenarios, as the top-side structure with short distances is not common.

With the rapid advancement of various new principles in solid-state and semi-solid-state LiDAR technologies [3], a type of LiDAR featuring a non-repetitive scanning pattern has gained widespread application [4]–[6]. They possess not only long sensing distances up to several hundred meters but also the capability to stack point cloud density in the time domain, thereby constructing dense geometric structures. These LiDARs guarantee the completion of the aforementioned top-

*This work was supported by the National Natural Science Foundation of China (U22A20100 / 62203294 / 62373250 / 62173228)

Xin'ao Wu, Chenxi Yang, Chunxiang Wang, and Ming Yang are with the Department of Automation, Shanghai Jiao Tong University, Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, 200240, China (email: mingyang@sjtu.edu.cn).

Yiyang Guo is with the Los Altos High School, Los Altos, CA, USA.

Hanyang Zhuang is with the University of Michigan - Shanghai Jiao Tong University Joint Institute. Shanghai Jiao Tong University, Shanghai, 200240, China.

side visibility, but their relatively narrow FoV for individuals remains an issue. After all, fully covering a hemispheric FoV requires many such LiDARs to be mounted on vehicle rooftops at different angles, which is not only uneconomical but also technically difficult.

In this study, we mount a single one of such LiDARs on a two-axis rotating gimbal, as shown in Fig. 1. Such a setting allows for the acquisition of an asynchronous hemispherical FoV with full coverage. We first utilize this device for incremental mapping, effectively supplementing the structural missing present at elevated positions in maps generated through traditional methods.

During the re-localization process, we leverage the structural information from the updated map to design automated gimbal motion strategies that ensure the LiDAR is consistently aligned with areas featuring geometric solid constraints within the environment. This approach facilitate the attainment of sufficiently precise pose estimations for the vehicle within a specific narrow FoV map matching condition. The experimental results validated the effectiveness of the proposed approach.

The main contributions are as follows:

(1) A LiDAR re-localization hardware configuration provides elevated geometrical constraints beyond the range of conventional perception;
(2) An active gimbal motion strategy to focus on the strong geometric constraint for map matching, which proves the feasibility of accurate small FoV pose registration.

II. Related Works

Following the discourse structure outlined in the last section, the literature review on re-localization is delineated into two primary aspects: utilizing LiDAR scans directly from perception and leveraging additional information from various signal sources.

A. Re-localization Sharing Perception's LiDAR Scans

As mentioned earlier, dynamic disturbances, in the context of re-localization tasks, serve not only as objects for detection but also as obstacles masking environmental structures. Numerous methods for detecting dynamic objects have been proposed [7]–[9]; our focus here is on discussing research related to matching residual static environmental structures.

The essence of maintaining robustness in map matching under occlusion scenarios lies in enhancing the efficiency of geometric constraints within local structures. Since the inception of the Iterative Closest Point (ICP) algorithm, many of its variants have aimed at enhancing such matching capabilities.

Since the inception of the Iterative Closest Point [10], a plethora of its' variants [11], [12] has aimed at enhancing such matching capabilities, as well as Normal-distributions Transform [13]. Among these efforts, LiDAR Odometry and Mapping (LOAM) [14] extracted edges and planes and achieved great performance in small FoV LiDAR mapping [4]. Semantics such as poles [15]–[17], road surface marks [18]–[20], and traffic lights [21] are often extracted independently or in combination [22] for such purposes. Unfortunately, no feature or object exists universally in every scenario, especially under severe occlusion of view.

B. Constructing Re-localization's Own Signal Sources

Given the limitations inherent in utilizing perception LiDAR scanning data, constructing independent signal sources represents another crucial approach for re-localization tasks.

Detecting magnetic nails buried beneath roads [23] or computing communication delays with several wireless base stations [24] can achieve precise re-localization with minimal interference. On this technological route, we have explored infrastructure-side sensor arrays for vehicle localization [25] through V2X communication. These implementations incur high costs and require time-consuming environmental modifications, making them suitable only for specific small-scale applications, such as parking lots or factory logistic systems.

Relatively lightweight solutions, such as adhering Apriltags [26] or reflective markers [27] to surfaces like ceilings with a camera to observe, offer a degree of flexibility in re-localization systems. However, their utility is still constrained to environments conducive to installing such auxiliaries.

Our last attempt to build a re-localization signal source without any infrastructure modification happened in factory workshops with a unique roof truss structure for LiDAR to generate abundant 3D structures [2]. As discussed in the last section, such customization makes it difficult to expand the use to wide-open areas due to the vulnerability of the hemispherical FoV LiDAR in forming valuable structures for remote environments. Meanwhile, such on-site deployments have provided crucial references for this study.

III. Methodology

This work aims to develop a top-view LiDAR re-localization method that can be implemented in arbitrary scenarios without environmental modifications. This section commences with an introduction to hardware.

A. Hardware Configuration

The non-repetitive scanning LiDARs, characterized by their unique optical structure, as shown in Fig. 2, are capable of temporally stacking point clouds to form dense structures, which makes them particularly suitable for this study. We selected the Livox Avia, featuring a FoV of around 70° and a maximum detection range of 450m, which offers advantages such as lightweight design and long sensing range. To address the common limitation of a small FoV in such LiDARs and thus the challenge of individually covering the top hemisphere FoV of the vehicle, we mounted a single Avia on a two-axis gimbal with a maximum angular speed of 35°/s. This configuration allows the LiDAR to asynchronously cover the top hemisphere through azimuth and elevation angles.

Since our test vehicle is equipped with a multi-layer LiDAR for comparison, the zero position of the gimbal relative to the vehicle's coordinate system, i.e., the yaw θ and pitch ϕ of the gimbal when the non-repetitive LiDAR is

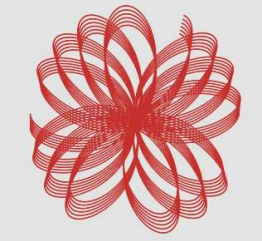


Fig. 2. The non-repetitive scanning circle of the selected Avia LiDAR.

horizontally oriented towards the vehicle's forward direction, can be simplified to a point-to-plane registration problem inter-LiDARs as:

$$(\theta^*, \phi^*) = \arg\min_{\theta,\phi} \sum_{i=1}^{k} \left(n_i^T (Rp_i + t^* - q_i)\right)^2 \quad (1)$$

where q_i is the point of the 3D multi-layer LiDAR, p_i is the point of Avia, and n_i is the normal vector of q_i. In this registration problem,

$$\begin{aligned} R &= R_z R_y R_x = R_z R_y \\ &= \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix} \end{aligned} \quad (2)$$

is the estimation target. Rolling angle R_x is ignored since the two LiDARs are mounted on the same plane and are close to each other. $t^* = [t_x, t_y, t_z]^T$ are fixed translation vector that is to obtain through manual measurement.

As demonstrated in Fig. 1(a), the Avia LiDAR is mounted at the center of the gimbal, thus rendering negligible any translations along the x and y axes relative to the two rotation centers of yawing θ and pitching ϕ. Furthermore, since the azimuthal axis of θ is orthogonal to the optical direction of the LiDAR, its relative z-axis translation can also be ignored. The coordinate transformation by the gimbal rotation needs to consider the z_{pitch} parameter between the rotation axis of ϕ and the optical center of the LiDAR as:

$$p' = R_z R_y (p + [0, 0, -z_{pitch}]^T) \quad (3)$$

After the joint registration of Equ. 1 and coordinate transformation of Equ. 3, the Avia LiDAR is registrated to the vehicle coordinate given any θ and ϕ.

B. Gimbal Motion Strategy

To accomplish re-localization through additional topside visibility as previously described, we devised a preparation plan that involves enhancing the map of that area to ensure map-side dense point cloud structures for matching. Considering the relatively consistent nature of building facades at various elevations, we opt to detect the wall structures of the lower elevation regions using a 2D grid map, which allowed us to determine the scanning areas for higher levels as follows:

for each point g_i in the grid map, its direction and salience can be represented as:

$$V(g_i) = \sum_{j=1}^{k} |g_{ij} - g_i| \quad (4)$$

where $g_{ij} \in N(g_i) = \{g_{i1}, g_{i2}, ..., g_{ik}\}$ is a k nearest neighbor of g_i found by k-d tree. While $\|V(g_i)\| = ({v_x}^2 + {v_y}^2)^{\frac{1}{2}}$ represents whether g_i could be a potential wall point, $[-v_y, v_x]$ determines the direction of the wall and guides the gimbal to enhance its high elevation part into the map. Fig. 3 shows a localized segment of the 2D grid map of the lower-elevation, identifying the facades and their orientations. By matching the building facades on the 2D map, we achieved attaching the elevation supplement map to the original 3D map.

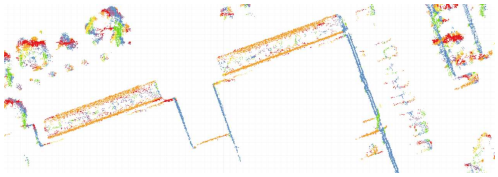


Fig. 3. The original map is first transformed into a 2D grid form for efficient facades detection, which is used to determine the gimbal motion strategy for high-elevation area map enhancement.

For scenes requiring enhancement in multiple directions within the environment, we conducted statistical analysis on the coverage duration of various structures in the grid map, considering that the density of point clouds from non-repetitive scanning LiDAR typically approaches its peak after stacking for around 3 seconds. This approach ensures a swift transition to the next segment of the facade after a segment has been illuminated.

The existing re-localization results from the multi-layer LiDAR can support the pose estimation during the point cloud stacking for map enhancement. However, the re-localization using our proposed equipment is targeted at scenarios with high occlusion that require the capability to achieve high-precision pose convergence independently under geometric constraints. Therefore, it's imperative to develop different gimbal motion strategies for the map-matching-based re-localization process.

Initially, we perform spatial segmentation on the 3D point cloud map's high-elevation regions through any potential yawing and pitching strategy (θ, ϕ). For each segment, we denote the mixture of the plane direction and segment scale as:

$$\mathbf{r}_i = (x_i, y_i, z_i)^T, \|\mathbf{r}_i\| = N_i \quad (5)$$

where $\mathbf{r}_i$ is the normal vector of plane segment i, with the modulus length of segment point number N_i. N_i represents the weight of each normal vector. Planes with more points exert greater influence on the orientation of planes within the segment.

Subsequently, we calculate the deviation of each weighted

normal vector from the mean as:

$$\begin{aligned}\delta \mathbf{r}_i &= (\delta x_i, \delta y_i, \delta z_i)^T \\ &= \mathbf{r}_i - \frac{1}{n}\sum_{i=1}^{n} \mathbf{r}_i \end{aligned} \tag{6}$$

The variance of the normal vector, including x and y components, is calculated as:

$$\Sigma_{xy}^2 = \sigma_x^2 + \sigma_y^2 = \frac{1}{n}\sum_{i=1}^{n}(\delta x_i^2 + \delta y_i^2) \tag{7}$$

Given any motion strategy (θ, ϕ), Σ_{xy}^2 represents the richness of the normal vectors' direction on the plane and the scale from point numbers. We track those strategies with minimum normal vector variance Σ_{xy}^2, which means the planes within these strategies have the best consistency. This study sets a 5° step for both axes and determines the gimbal motion strategy for the re-localization stage.

IV. Experiments

Fig. 4 illustrates our experimental setup. To prevent the proposed equipment from obstructing the baseline LiDAR, we elevated the baseline LiDAR to a greater height. While ensuring the complete FoV of the baseline method, the FoV of our proposed system will inevitably be partially blocked by the baseline. We masked this part of the FoV in the gimbal rotation strategy. We used a Velodyne VLP-16 as the baseline of the conventional 3D multi-layer LiDAR, another multi-layer LiDAR Hesai Pandar40 for comparison, and the ground truth is provided by a high precision RTK-GPS-IMU system, which has an accuracy of 1cm.

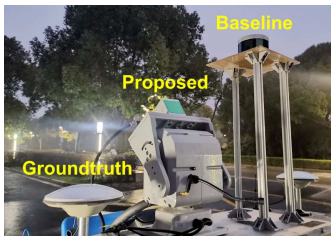


Fig. 4. The experimental setup. For fairness, we sacrificed a part of our FoV to ensure the full FoV of the baseline system.

We first placed the equipment indoors and reconstructed the environment based on the gimbal's angular feedback to compare the FoV between one frame and the asynchronous complete FoV coverage, as seen in Fig. 5. Fig. 6(a) demonstrates our experiment scene, Route 1, which is a 650m long main road in our campus. Along the road there are buildings with distinct facades. We also tested our method in an extreme scenario, Route 2, as shown in Fig. 6(b). It is a 310m long route with facades only at a distance.

The vehicle is manually driven at the speed of 15 km/h, simulating relatively low-speed scenarios typically of schools, parks, and similar environments. The gimbal and the upper computer are connected via Ethernet cable, through which the upper computer controls the gimbal rotation or reads the angle by sending control commands. Each control command is a series of hexadecimal codes. The gimbal rotation speed is set at 22°/s.

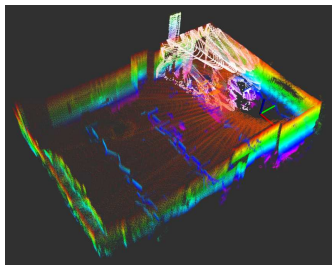
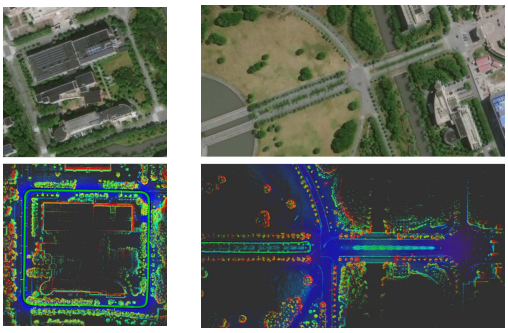
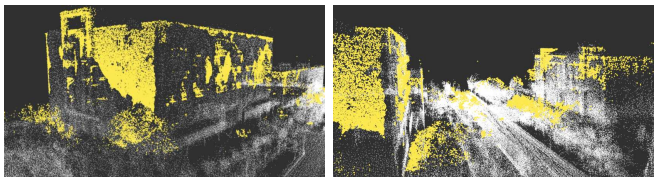


Fig. 5. FoV comparison. Indoor environemnt (colored) reconstructed by a series of single frames (white).



(a) (b)

Fig. 6. The experimental sites. The map is a combination of surveying based on high-precision GPS pose and our map enhancement method. (a) Main road with rich facades. (b) Areas with few buildings. The only facades advisable is at a distance.



(a) (b)

Fig. 7. Map enhancement at high-elevation. White: original re-localization oriented map generated by 3D multi-layer LiDAR. Yellow: completed part of the building facades.

Fig. 7 demonstrates the map enhancement result via a one-way data collection that corresponds to the first part of Section III-B. Compared to the original environment map, the re-localization generated by the conventional 3D multi-layer LiDAR displays the facades of the buildings with sharp edges.

On Route 1, we conducted one experiment during rush hours and another in the early morning to compare the performance of the proposed method under highly obstructed and unobstructed environments. Initially, we evaluated the improvement in the re-localization failure rate during severe dynamic occlusion in rush hours. We set the initial pose for point cloud registration from the last frame and

evaluated the registration shift in translation and rotation. Translations larger than 0.3m or heading shifts greater than 3° are regarded as a failure and a potential risk for future pose registration. If this failure occurs, we re-initialized the pose for the next frame and kept the calculation using the GNSS-based ground truth. As shown in Table I, our proposed method reduced such potential risk by about two times compared to multi-layer LiDARs, while there was only a marginal decrease in failure rate from 16-layer to 40-layer LiDAR. This demonstrates that ensuring robust features for matching is a more important factor for re-localization and proves the solution's effectiveness has met our design purpose.

Table II demonstrates the performance of our proposed system without any dynamic disturbance. We also provided an ablation concerning using the Avia LiDAR without the gimbal rotation (Avia w/o rot).

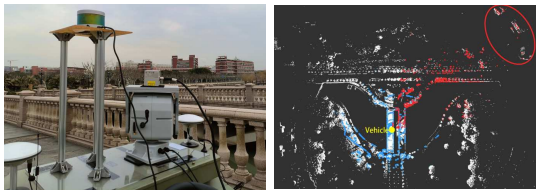
TABLE I

PERFORMANCE COMPARISON ON ROUTE 1 UNDER SEVERE DYNAMIC INTERFERENCE CONDITIONS AMONG DIFFERENT HARDWARE SETTINGS

Method	Re-localization failure rate (%)
Multi-layer(16)	17.50
Multi-layer(40)	13.40
Ours	**6.25**

In this experiment, we observed that the forward-facing small FoV LiDAR failed to achieve effective re-localization, even in an environment that does not contain any dynamic interferences. This result was evident in the validation, where we halted the verification midway upon the occurrence of degradation in point registration performance, as indicated by the (*) in Table II. This failure serves as a lateral validation of the rationale behind our employment of the gimbal system.

By comparing this performance with that of the multi-layer LiDAR, we observed that in the ideal unobstructed conditions, the pose constraint capability of a small FoV needs further effort to cover 360°. However, from a practical standpoint, the Avia LiDAR already provides a valuable reference for path planning in many autonomous driving applications.

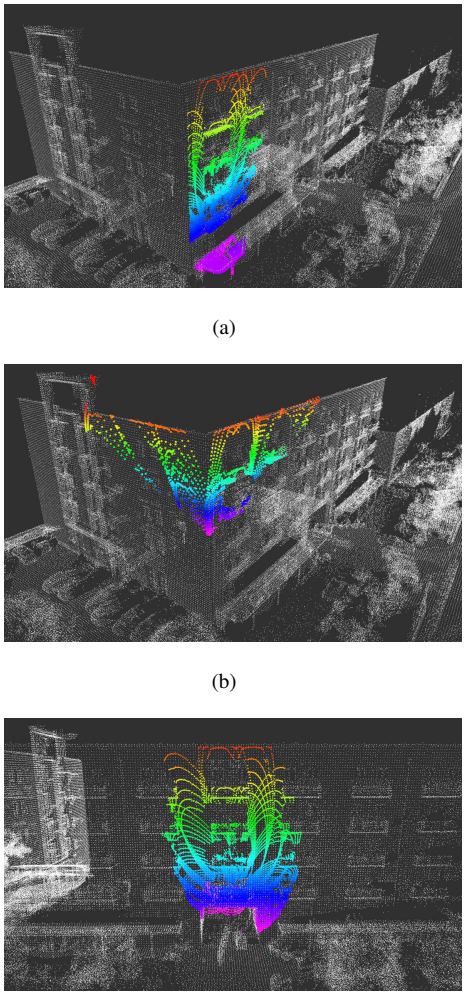


(a) (b)

Fig. 8. Validation of effectiveness conducted in under-structured areas. (a) Scenario. (b) The motion strategy determined remote building facades as the re-localization target, which demonstrated the capability of the proposed system to deal with potential structural degradation. Blue: multi-layer LiDAR point cloud. Red: proposed.

For Route 2, Fig. 8 briefly demonstrates the adaptability of our system. In under-structured areas that have relatively low dynamic environmental pressure, our proposed gimbal motion strategy still directs the Avia LiDAR to face the remote facades up to 350m, resulting in successful re-localization thanks to low structures in between. Under the same condition, the multi-layer LiDAR only has a scan range within 150m, resulting in potential insufficient structures for matching. In the event of severe obstruction occurring in such scenarios, the proposed method can still serve as a re-localization option.

More experimental results are demonstrated in Fig. 9.



(a)

(b)

(c)

Fig. 9. Point registration results at different poses. (a) With a significant pitching angle, the FoV is unlikely to be entirely obstructed by dynamic objects. However, awnings pose a potential challenge even for the proposed system. It's advisable to avoid setting routes under such structures, if possible, to prevent potential degradation. (b) The gimbal motion strategy opted to rotate back before the vehicle turned to prevent a shortage of features for matching. (c) Facades sub-structure is sufficent to provide strong geometric constraints.

V. CONCLUSIONS

This study proposes a gimbal-mounted active LiDAR re-localization system that utilizes a two-axis gimbal. Through techniques such as high-altitude map completion and gimbal

TABLE II

PERFORMANCE COMPARISON ON ROUTE 1 UNDER NO INTERFERENCE CONDITION AMONG DIFFERENT HARDWARE SETTINGS

Method	Translation error RMS (m)	Translation variance	Rotation error RMS(°)	Rotation variance
Multi-layer(16)	0.11	0.015	0.11	0.013
Multi-layer(40)	**0.10**	**0.008**	**0.09**	**0.007**
Avia w/o rot	1.39*	12.93	-	-
Ours	0.21	0.060	0.35	0.16

motion strategy design, the system enables a small FoV non-repetitive scanning LiDAR to asynchronously cover the upper hemisphere FoV on an intelligent vehicle. The proposed method surpasses traditional multi-layer LiDARs in scanning distant building facades, thereby exhibiting enhanced re-localization capabilities in scenarios characterized by severe dynamic occlusion, such as near by buses.

Through experiments, we validated that the proposed method not only enhances performance when combined with traditional implementations but also can achieve reliable re-localization in typical scenarios independently. This provides an alternative with certain universality in this field. In the future, we would focus on lightweight gimbal equipment and optimizing gimbal motion strategies to better adapt to challenges posed by various scenarios.

REFERENCES

[1] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann *et al.*, "Stanley: The robot that won the darpa grand challenge," *Journal of field Robotics*, vol. 23, no. 9, pp. 661–692, 2006.

[2] Y. Wu, W. Yuan, B. Wang, C. Lian, Y. Yao, Y. Cai, and M. Yang, "Truss feature based robust localization method for vehicles in dynamic industrial scene," in *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2023, pp. 1–6.

[3] D. Van Nam and K. Gon-Woo, "Solid-state lidar based-slam: A concise review and application," in *2021 IEEE International Conference on Big Data and Smart Computing (BigComp)*. IEEE, 2021, pp. 302–305.

[4] J. Lin and F. Zhang, "Loam livox: A fast, robust, high-precision lidar odometry and mapping package for lidars of small fov," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 3126–3131.

[5] K. Li, M. Li, and U. D. Hanebeck, "Towards high-performance solid-state-lidar-inertial odometry and mapping," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5167–5174, 2021.

[6] T. Hu, X. Sun, Y. Su, H. Guan, Q. Sun, M. Kelly, and Q. Guo, "Development and performance evaluation of a very low-cost uav-lidar system for forestry applications," *Remote Sensing*, vol. 13, no. 1, p. 77, 2020.

[7] A. Petrovskaya and S. Thrun, "Model based vehicle detection and tracking for autonomous urban driving," *Autonomous Robots*, vol. 26, no. 2-3, pp. 123–139, 2009.

[8] A. Dewan, T. Caselitz, G. D. Tipaldi, and W. Burgard, "Motion-based detection and tracking in 3d lidar scans," in *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2016, pp. 4508–4513.

[9] B. Ravi Kiran, L. Roldao, B. Irastorza, R. Verastegui, S. Suss, S. Yogamani, V. Talpaert, A. Lepoutre, and G. Trehard, "Real-time dynamic object detection for autonomous driving using prior 3d-maps," in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018, pp. 0–0.

[10] P. BESL and N. MCKAY, "A method for registration of 3-d shapes," *IEEE transactions on pattern analysis and machine intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[11] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp." in *Robotics: science and systems*, vol. 2, no. 4. Seattle, WA, 2009, p. 435.

[12] J. Serafin and G. Grisetti, "Nicp: Dense normal based point cloud registration," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015, pp. 742–749.

[13] M. Magnusson, "The three-dimensional normal-distributions transform: an efficient representation for registration, surface analysis, and loop detection," Ph.D. dissertation, Örebro universitet, Stockholm, 2009.

[14] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Autonomous Robots*, vol. 41, no. 2, pp. 401–416, 2017.

[15] R. Spangenberg, D. Goehring, and R. Rojas, "Pole-based localization for autonomous vehicles in urban scenarios," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Daejeon, Korea: IEEE, 2016, pp. 2161–2166.

[16] M. Sefati, M. Daum, B. Sondermann, K. D. Kreisköther, and A. Kampker, "Improving vehicle localization using semantic and pole-like landmarks," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. Redondo Beach, CA, United states: IEEE, 2017, pp. 13–19.

[17] A. Kampker, J. Hatzenbuehler, L. Klein, M. Sefati, K. D. Kreiskoether, and D. Gert, "Concept study for vehicle self-localization using neural networks for detection of pole-like landmarks," in *International Conference on Intelligent Autonomous Systems*. Baden-Baden, Germany: Springer, 2018, pp. 689–705.

[18] F. Poggenhans, N. Salscheider, and C. Stiller, "Precise localization in high-definition road maps for urban regions," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Madrid, Spain: IEEE, 10 2018, pp. 2167–2174.

[19] A. Hata and D. Wolf, "Road marking detection using lidar reflective intensity data and its application to vehicle localization," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. Qingdao, China: IEEE, 2014, pp. 584–589.

[20] J. K. Suhr, J. Jang, D. Min, and H. G. Jung, "Sensor fusion-based low-cost vehicle localization system for complex urban environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1078–1086, 2016.

[21] C. Wang, H. Huang, Y. Ji, B. Wang, and M. Yang, "Vehicle localization at an intersection using a traffic light map," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1432–1441, 2018.

[22] C. Yang, L. He, H. Zhuang, C. Wang, and M. Yang, "Pseudo-anchors: Robust semantic features for lidar mapping in highly dynamic scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 2, pp. 1619–1630, 2023.

[23] G. Zhu, M. Yang, H. Li, B. Wang, and C. Wang, "Curvature map-based magnetic guidance for automated vehicles in an urban environment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 12, pp. 3541–3552, 2016.

[24] H. Wang, L. Wan, M. Dong, K. Ota, and X. Wang, "Assistant vehicle localization based on three collaborative base stations via sbl-based robust doa estimation," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5766–5777, 2019.

[25] F. Wang, Y. He, H. Zhuang, C. Yang, and M. Yang, "Cross-modal registration using adaptive modeling in infrastructure-based vehicle localization," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. Yokohama, Japan: IEEE, 2024, pp. 18 377–18 383.

[26] E. Olson, "Apriltag: A robust and flexible visual fiducial system," in *2011 IEEE international conference on robotics and automation*. IEEE, 2011, pp. 3400–3407.

[27] R. Muñoz-Salinas, M. J. Marín-Jimenez, E. Yeguas-Bolivar, and R. Medina-Carnicer, "Mapping and localization from planar markers," *Pattern Recognition*, vol. 73, pp. 158–171, 2018.