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

PS-Loc: Robust LiDAR Localization with Prior Structural Reference

Rui Li, Wentao Zhao, Tianchen Deng, Yanbo Wang and Jingchuan Wang*, *Senior Member, IEEE*

***Abstract*— Prior structural reference like floor plan is readily accessible in indoor scene, which exhibits the potential of improving localization quality without the requirements of a previously-built high-precision map. This paper introduces a novel optimal transport-based framework for prior structural reference-based localization, aiming to improve the robustness for the robot localization. Leveraging the spacial relations of structures, a matching method based on optimal transport theory is proposed and it improves the robustness of matching results in dynamic scene and rapid rotation conditions. Additionally, this paper handles metric inaccuracies in the known structural reference by implementing an prior guided plane adjustment-based updating strategy. This strategy combines prior and observational information to jointly optimize the structural information within a sliding window. The performance of the framework is validated through real-world experiments, demonstrating superior accuracy and robustness to disturbances from dynamic occlusion and rapid rotation compared to common state-of-the-art SLAM and localization methods.**

***Index Terms*— LiDAR Localization, Prior structural reference, Optimal transport-based matching**

I. Introduction

LiDAR localization stands as a fundamental technique for autonomous mobile robotics, typically relies on maps generated through simultaneous localization and mapping (SLAM) techniques [1]–[3]. However, constructing a high-precision map via SLAM for LiDAR localization can be a laborious and time-consuming process, especially in dynamic environments. The presence of moving objects within the environment can introduce ghost artifacts into the maps, directly compromising the robustness and precision of localization. There is a growing trend of assisting LiDAR localization and mapping with prior knowledge [4]–[6] to improve localization and mapping quality and alleviate repetitive work of high-precision map creation before tasks.

The most commonly used and accessible source of prior information is architectural structural references generated during building construction, such as floor plan sketches, wall segmentation, CAD drawings and so on. Several works **regard the structural priority as existed high-precision landmarks** to improve localization accuracy. Boniardi et al. [7] propose a pose graph based localization method with constraints between point cloud and CAD floor plans by Generalized ICP(GICP). This approach improves the accuracy and global consistency of 2D pose estimation. Furthermore, the framework is extended [4] for long-term localization by detecting dynamic point clouds. A similar approach was adopted by Mielle et al. [8] through considering scale inaccuracy. Oelsch et al. [5], [9] extract 3D triangular mesh features from 3D models and constraint the points near the model by point-to-mesh residuals, which greatly reduces the absolute pose error(APE) in simulation. However, these methods rely heavily on precise prior information, and any imperfections can have negative effect on localization. And simply employing ICP-based searching methods makes matching sensitive to initial pose and less robust to rotation. Several other works carefully design methods to **convert prior information into another data representations**, aiming to enhance the system's robustness and efficiency. Xu et al. [10] generate scans using a prior 3D map and achieve robust place recognition with virtual point cloud. Gao et al. [11] propose a data structure that enables points to efficiently look-up the corresponding geometric floor plan elements. But disturbance like changing cabinets also introduce geometric information and interfere with matching results. Luo et al. [12] design architectural feature patterns and descriptors based on the Manhattan world assumption, using architectural skeletons extracted from prior reference for coarse registration, reducing drift in non-loop areas.

Architectural structural references provide stable building information and their overall relationships like relative positions. However, leveraging such information remains several challenges. **1) Inaccurate structural reference.** Although structural reference can provide globally consistent constraints, it can not guarantee local metric accuracy for sketch of floor plan. The environments inevitably changes in practice, and simply aligning point cloud to prior impairs localization under such situation, necessitating consideration for updating them. **2) Dynamic occlusion.** Although structural reference only contains 2D static information, dynamics in 3D space can cause occlusion and even introduce additional dynamic structural features, potentially leading to matching confusion. **3) Rapid rotation.** LiDAR-based pose estimation yields poorer results in rotation compared to translation, as long range points moves a lot with large radius, and poses a poor initial guess for nearest searching methods. In order to robustly search for matching from a more perspective way, the matching method leveraging the relation of structural features needs exploring.

To address these issues, we introduce a robust prior structural reference-based localization framework called PS-Loc. Inspired by efficent and effective geometrical SLAM [13]–[15] , our method first extracts structural features from

This work was supported by the National Natural Science Foundation of China under Grant U2013203. The authors are with Institute of Medical Robotics and Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education, Shanghai 200240, China. (E-mail: li_rui@sjtu.edu.cn; wentaozhao@sjtu.edu.cn; dengtianchen@sjtu.edu.cn; yanbowang319@sjtu.edu.cn; *corresponding author: jchwang@sjtu.edu.cn;)

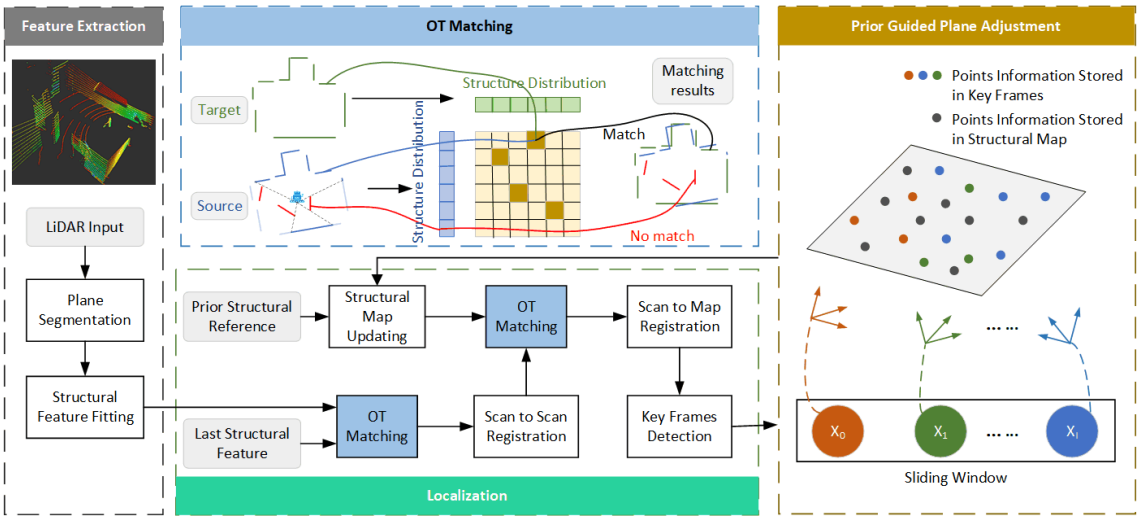


Fig. 1: System overview.

LiDAR scans. Then, to handle matching confusion caused by dynamics and rapid rotation, an optimal transport-based matching (OT matching) method is proposed. Finally, after registering with the structural map, the metric inaccuracy of structural map is updated using the prior guided plane adjustment (PG-PA) method [16], [17] to enhance localization precision. More specifically, the main contributions of this work can be summarized as follows:

- A novel matching method is proposed based on optimal transport theory to establish correspondences between the structural features from the prior map and those from the current observations. This approach effectively addresses disturbances arising from rapid ego-motion, dynamic occlusions and disparities between observational data and prior information.
- An incremental updating strategy combining prior information and observations is employed to address the incomplete local information provided by the structural reference and enhance localization precision.
- Extensive experiments are conducted in various scenarios. The results validate the system's high precision and robustness outperforming the state-of-the-art even when faced with disturbances caused by rapid ego-motion and dynamic objects.

II. Methodology

The structure of proposed LiDAR localization method is shown in Fig. 1. The system starts with feature extraction by plane segmentation followed by structural feature fitting (see Subsection II-A). This process links 3D LiDAR points with 2D structures. Then the optimal transport-based matching (OT Matching) algorithm (see Subsection II-B) establishes robust correspondences between prior and observational structural features, even in the presence of dynamics and rapid rotation interference. After registering with the global map, key frames are selected based on translation and rotation criteria. Then, the PG-PA in a sliding window and map updating strategy (see Subsection II-C) are adopted to update metric information of the structural map and enhance localization precision.

A. Structural Feature Extraction

Priority provides properties of structures, such as the location, direction and length. This subsection discusses structural feature extraction from 3D LiDAR scans, and parameters stored for later registration with little information loss.

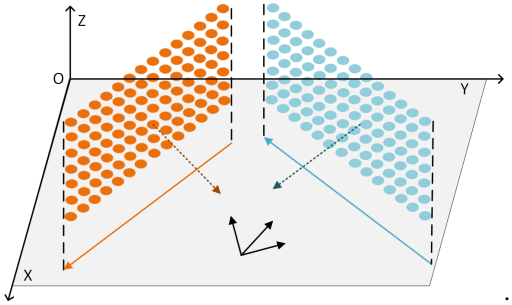


Fig. 2: Structure fitting from 3D LiDAR points.

The prior structure here is planar building information. Given the importance of real-time performance for mobile robots, a fast multi-plane detection method is required [18]. Projected onto a range image, points can be clustered into lines and then grouped into planes using a breadth-first searching method. This hierarchical segmentation approach reduces time consumption compared to brute-force searching point by point. To fit structural features from the detected planes, the two farthest boundary points of each vertical plane segments are first selected, and the time complexity is only O(N) to find the max distance from each point to cloud center in two directions. Those boundary points are projected onto the fitted plane, then onto the ground, forming a structural feature composed of starting and ending points, see Fig. 2. The direction of a structural feature is determined by rotating the plane's normal vector, pointing at the robot,

by -90 degrees around the z-axis. This distinguishes the two sides of each structural wall. When fitting structures, all plane points are stored in a parameter A without planar information loss,

$$A_j = \sum_{k=0}^{K_j} \bar{p}_{jk}\bar{p}_{jk}^T = \begin{bmatrix} \sum_{k=0}^{K_j} p_{jk}p_{jk}^T & K_j p_{jc} \\ K_j p_{jc}^T & K_j \end{bmatrix}, \quad (1)$$

where $\bar{p}_{jk} = [p_{jk}^T; 1]^T$ is the homogeneous coordinates of kth point in jth plane, A_j contains information of cloud size K_j, center p_{jc} and covariance $Cov_j = \sum_{k=0}^{K_j} p_{jk}p_{jk}^T/K_j - p_{jc}p_{jc}^T$.

B. Structural Feature Matching Based on Optimal Transport Theory

After fitting 2D structural vectors, the association problem between point clouds and structural references turns into matching problem between two structural feature sets.

1) Problem Formulation: The classical formulation of static Kantorovich formulation [19] is show in Equation (2). For two sets with the same total mass α, β on a set X and with a transport effort $c(x, y) \longmapsto \mathbb{R} \cup \infty$, the assignment problem between α and β is,

$$inf \left\{ \int_{X^2} c(x,y)d\gamma(x,y), (Proj_0)_{\#\gamma} = \alpha, (Proj_1)_{\#\gamma} = \beta \right\}, \quad (2)$$

where $c(x, y)$ is the transport cost from x to the y, $\gamma(x, y)$ is a non-negative transport plan on X^2 that satisfies the marginal constraints, and $(Proj_m)_{\#X^2}$ denotes the canonical projection from X^2 onto factor m.

Optimal transport theory aims to transport mass from one set to the other within minimal effort considering distribution of two sets, and the mass here is inherently the length of each structural features. The matching results between structural feature sets can be rounded off from the transport matrix with larger values, see Fig. 1. However, in practice, environmental dynamics cause occlusion and add features, making differences in total mass between observational and prior structures. To address this situation, bins with a fixed transport cost are added to each set to equalize the total mass, and features with high transport losses are eliminated by matching bins. The discrete formulation of such problem are defined as below:

$$\begin{aligned} OT(\alpha,\beta) = & \min_{\gamma_{i,j} \in \mathbb{R}_{\geq 0}^{N \times M}} \sum_{i=1}^{N}\sum_{j=1}^{M} \gamma_{i,j} c(x_i, y_j) - \varepsilon H(\gamma_{i,j}) \\ s.t. & \sum_{j=1}^{M} \gamma_{i,j} = \alpha, \sum_{i=1}^{N} \gamma_{i,j} = \beta, \end{aligned} \quad (3)$$

where α and β are the length distribution of surrounding structural features with miss matching bins added, $c(x_i, y_j) = \|\vec{d}(x_i, y_j)\|_2$ represents the moving effort from structure x_i to structure y_j. The entropy term $H(\gamma_{i,j}) = -\sum_{i=1}^{N}\sum_{j=1}^{M} \gamma_{i,j} \log(\gamma_{i,j} - 1)$ regulates this problem strongly convex, so this problem can be solved by Sinkhorn algorithm.

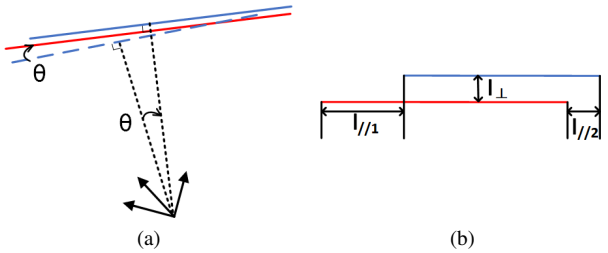


Fig. 3: Two steps transporting source line to the target line. (a) Step 1: Rotation displacement. (b) Step 2: Translation displacement.

Optimal transport-based structural feature association method aligns the overall distribution of observance with the priorities. This eliminates local disturbances caused by dynamic occlusion and is less sensitive to the initial pose than nearest search-based methods [20] like rapid rotation.

2) Structure Displacement Cost: Inspired by [21] the transport effort vector $\vec{d}(x_i, y_j)$ is defined as:

$$\vec{d}(x_i, y_j) = \begin{bmatrix} f(\theta(x_i, y_j)) \\ l_\perp(x_i, y_j) \\ \min(l_\parallel(x_i, y_j), l_\parallel(y_j, x_i)) \end{bmatrix}, \quad (4)$$

where $f(\cdot)$ is a penalty function mapping angular difference θ between two structures to a scalar, see Fig. 3a.The dashed blue and solid red lines represent source and target structure respectively. Rotate the source structure θ to align it with the target, see Fig. 3b. Then, align them costs $l_\perp(x_i, y_j)$ in vertical direction and $l_\parallel$ in parallel direction, which the non negative $l_\parallel(x_i, y_j)$ means the length x_i exceeds y_j.

As detailed in [22], Sinkhorn solver returns a pair of dual vectors $f \in \mathbb{R}^N$ and $g \in \mathbb{R}^M$ that encode the transport plan by

$$\gamma_{i,j} = exp\left(-\frac{1}{\varepsilon}(c(x_i, y_j) + f_i + g_j)\right). \quad (5)$$

Since the coupling is many-to-many, the largest value in matrix is selected as the matching results for each source structure.

C. Prior Guided Plane Adjustment

The OT Matching offers robust association results for point-to-plane error alignment and structural updating. A plane can be represented as a four-dimension vector $\pi = [\boldsymbol{n}^T; d]$ combined as a unit norm vector $\boldsymbol{n}$ and the distance from origin to this plane. Vector π can be derived from the plane parameter A, as $\boldsymbol{n}$ can be obtained from the minimum eigenvector of the point cloud covariance Cov and $d = -\boldsymbol{n}^T p_c$, see Equation (1).

1) Point-to-Plane Re-projection Error: Given correspondences, pose estimation can be achieved by minimizing the least square problem of point-to-plane re-projection error:

$$\min_T \sum_{j=0}^{J}\sum_{k=0}^{K_j} \|\pi_{\phi(j)} T p_{jk}\|^2 = \min_T \sum_{j=0}^{J} \pi_{\phi(j)} T A_j T^T \pi_{\phi(j)}^T, \quad (6)$$

where matrix T transfers source planes to the target coordinate, J is the number of planes detected, $\phi(j)$ is a permutation function mapping jth plane to its corresponding target plane which is solved in Subsection II-B. $\pi_{\phi(j)}$ represents the corresponding target plane of the jth source plane. And Scan-to-scan and scan-to-map registration in Fig. 1 are accomplished by minimizing this re-projection error.

2) Plane Adjustment Combining Prior Structures: While the structural relations from the prior reference can enhance localization robustness, inaccuracies in the metrics might compromise localization accuracy. A joint optimization in poses and structural features is adopted combining the information from prior reference with observations in a sliding window is put forward.

Given groups of plane point clouds from I key frames corresponding to the mth structural reference, the key frame poses $T_i, i = 1, 2, \ldots, I$, and the m th plane feature π_m are jointly determined by minimizing the sum of squared distances, mentioned in Equation (6), from each plane point in key frames to the structural reference. Note that the m th prior structure also includes parameter A_m, with its initial value derived from simulated scan points. Each structure represents a rectangular wall surface with a height of h meters. Points are uniformly sampled horizontally and vertically, with Gaussian sampling in the perpendicular direction. And then, the joint optimization with all M structural references is defined as follows:

$$\begin{aligned}
&\min_{\pi,x,y,yaw} \sum_{m=0}^{M} \left(\sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{k=0}^{K} \mu \|\pi T_i p_{ijk}\|^2 + \sum_{k'=0}^{K'} \|\pi p_{mk'}\|^2 \right) \\
&= \min_{x,y,yaw} \sum_{m=0}^{M} \underbrace{\left\{ \min_{\pi} \pi \left(\mu \sum_{i=0}^{I} \sum_{j=0}^{J} T_i A_{ij} T_i^T + A_m \right) \pi^T \right\}}_{=\lambda_1(Cov);\ \pi^*=[\boldsymbol{v_1};-\boldsymbol{v_1 p_c}]},
\end{aligned} \tag{7}$$

where $J = \phi^{-1}(m)$ is the number of planes related to the mth target structure. K and K' represent the number of points in observance and priority. T_i is the pose transformation matrix of the ith frame. μ gives a weight to prior and observation. Cov and p_c are the covariance and center of plane point cloud that can be derived from joint parameter $A = \mu \sum_{i=0}^{I} \sum_{j=0}^{J} T_i A_{ij} T_i^T + A_m$, and $\lambda_1(Cov)$ and v_1 are the minimal eigenvalue of covariance Cov and its corresponding eigenvector. Derivatives can be calculated following [16].

To address the in accuracy of prior structural reference in metrics, an incremental updating strategy are employed:

$$A_m \leftarrow A_m + \mu \sum_{i=0}^{I} \sum_{j=0}^{J} T_i A_{ij} T_i^T, \tag{8}$$

where the norm vector and center point will also be updated with gradual observations. New structural features and tilted planes are added if a detected plane, transport through a sliding window, without matching existing structures. This allows reliable 6 degree of freedom (DOF) pose estimation even with limited structures available by priority.

III. Experiments

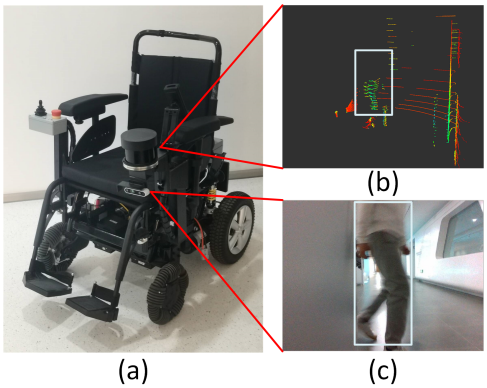


Fig. 4: Illustration of robot platform and data collection. (a) Picture of the intelligent wheelchair. The LiDAR has an 80-degree obscured field of view and cannot scan backward. The camera is right below the LiDAR. (b) Example of LiDAR point cloud when a pedestrian passing by. The entire left side of point cloud is disturbed, impairing localization.(c) Example of camera image with the same pedestrian marked by rectangle.

Real-world experimental data is collected using an intelligent wheelchair platform [23], see Fig. 4. Prior structures are derived from floor plans with a measurement error of approximately 5%. All experiments are calculated on a desktop equipped with an Intel i7-8700 CPU and 15GB of memory, achieving real-time performance. The ground truth is obtained using a well-calibrated motion capture system with millimeter-level positioning accuracy. The proposed method is a framework of leveraging and managing prior structural reference, facilitating integration into other SLAM systems. It is integrated into A-LOAM [1] by combining the loss terms to work even in scenes without sufficient planar information.

The comparison includes the open-source LiDAR-based SLAM system A-LOAM and its loop closure variant SC-A-LOAM. Additionally, a robust LiDAR-inertia localization method ROLL [24], based on FAST-LIO2 [25] and using well-calibrated IMU measurements from the RealSense D435i, is included for comparison. The map for ROLL is built within the same day and scene with no dynamics.

A. Experiments of Indoor Localization Precision

To evaluate the localization precision performance of the proposed method in narrow space, sequence 1 is collected in a square room with clutters. Table I lists the absolute trajectory error (ATE) results of each method. The suffix (2D) indicates the ATE of trajectory without the height estimation. In sequence 1, PS-Loc's mean and median APE in 6 DOF are under 2 cm, outperforming other methods in all metrics.

To evaluate the localization precision in a medium sized indoor space, sequence 2 is collected. It covers an area of 3000 m^2, including cluttered rooms, narrow corridors and halls. Ground truth poses are only available at the head

TABLE I: Drift Comparison on RS-LiDAR-16 Dataset

Method	Loop	IMU	sequence 1(static small)			sequence 2(static medium)			sequence 3(dynamic medium)		
			rmse	mean	median	rmse	mean	median	rmse	mean	median
SC-A-LOAM	✔	✘	8.142	9.606	7.775	57.399	55.336	52.058	150.658	119.099	69.556
A-LOAM	✘	✘	2.931	2.70	2.75	120.547	111.580	94.67	138.663	133.434	142.812
ROLL	✘	✔	2.519	2.224	2.057	3.357	2.350	**1.235**	**4.118**	3.272	3.272
PS-Loc	✘	✘	**2.292**	**1.918**	**1.788**	**2.323**	**2.029**	1.554	4.322	**2.990**	**1.022**
SC-A-LOAM(2D)	✔	✘	8.086	7.543	7.678	57.057	54.980	52.122	93.636	75.042	47.416
A-LOAM(2D)	✘	✘	2.450	2.110	1.998	3.024	2.745	2.391	49.403	46.477	44.970
ROLL(2D)	✘	✔	2.149	1.796	1.629	3.280	2.218	**1.035**	**4.036**	3.209	1.562
PS-Loc(2D)	✘	✘	**2.137**	**1.739**	**1.609**	**2.210**	**1.915**	1.435	4.237	**2.884**	**0.954**

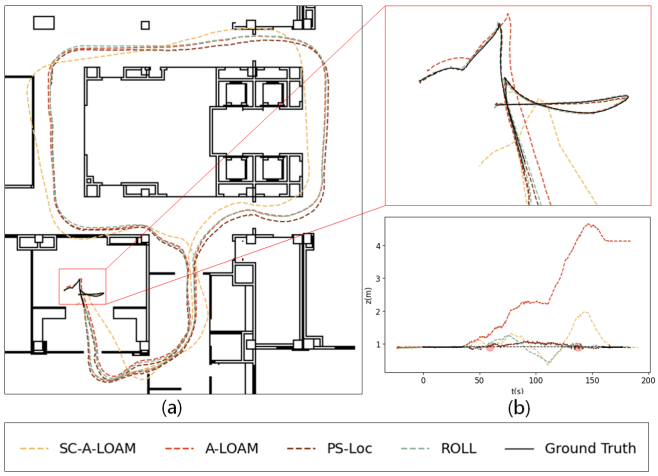


Fig. 5: Results from sequence 2. (a) Trajectory aligned with structural reference. (b) Height estimation comparison.

and tail ends, and the ATE of these two parts is used to represent repeated localization accuracy. In the narrow corridor scene, A-LOAM and SC-A-LOAM perform poorly in height estimation especially in high occlusion situations.

Although the loop closure of SC-A-LOAM reduces the overall APE compared to A-LOAM in Table I, and loops back at the 170s in height estimation, see Fig. 5b., the graph pose based loop closure disperses the overall error to each pose, resulting in a loss of local accuracy. Its trajectory obviously deviates from other trajectories in Fig. 5a. PS-Loc can achieve local high precision and global consistency thanks to the prior reference and local map updating strategy, and outperforms A-LOAM even not considering height drift. The height of robot changes at 59s and 138s, marked with red circles, as the ground plane's height varies when crossing a slope, see Fig. 5b. Only PS-Loc detect such height movement, although ROLL reduces height drift with the help of IMU.

B. Experiments in Dynamic Scene

To test the robustness to environmental dynamic interference, sequence 3 is collected. This data follows the same path as sequence 2, but with two people walking around and interfering with the LiDAR. The occlusion and dynamic features impair localization precision with higher APE compared to sequence 2. Furthermore, the performance of loop closure is damaged by dynamics in key frames and large height drift. Since the IMU used in ROLL is essentially unaffected by dynamics, ROLL's APE increases by only 1 cm. PS-Loc considers the overall distribution of surrounding structures, lowering the impact of local distribution differences caused by dynamics. And It outperforms all listed methods in mean and median APE.

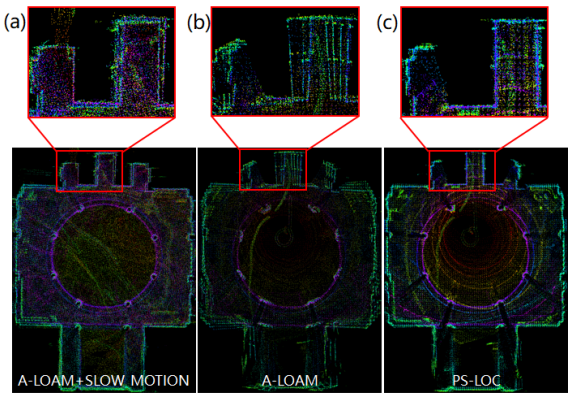
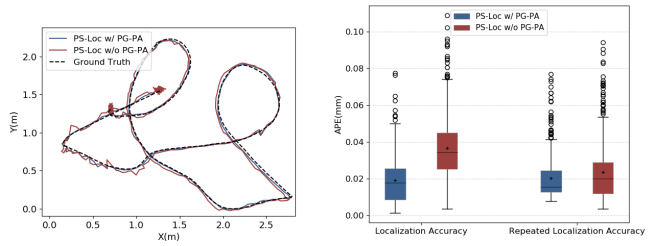


Fig. 6: Comparison of mapping results. (a) Reference map constructed by A-LOAM in the hall under slow motion (b) Map constructed by A-LOAM under rapid rotation. (c) Map constructed by PS-Loc under rapid rotation.

C. Experiments under Rapid Rotation

To evaluate the robustness of matching method during rapid rotation, sequence 4 is collected in a hall setting, where the robot's motion is more unrestricted. As PS-Loc is integrated to A-LOAM, they share the same point cloud map construction method. So the mapping results are used to evaluate the robustness of localization to rapid rotation (up to 30 degrees per second), see Fig. 6. Additional slow-motion data is collected for A-LOAM to build a benchmark map, serving as the standard for testing localization quality during rapid rotations. Rapid rotation results in significant movement of long-rang points, leading to a poor pose initial guess for cloud registration. ICP-based method falls into local minima and cannot compensate for errors by iteration, while optimal transport based structural matching method considers the distribution and relation of surrounding structural features and remains robustness even if point cloud moves a lot. Results show that PS-Loc exhibits robustness to rapid rotation and its map is closer to the one in slow motion condition. But A-LOAM undergoes trajectory drift, resulting in a blurred map.



(a) (b)

Fig. 7: Tests on plane adjustment and map updating. (a) Trajectory comparison in sequence 1. (b) Comparison of box plots for APE. Localization accuracy in smaller space and repeated localization accuracy in medium sized space are both compared.

D. Ablation Experiments

To assess the effectiveness of PG-PA, a comparative experiment is conducted between systems that include and exclude this module. The trajectory is shown in Fig. 7a, and the one adding this method is smoother. The map updating method evidently improves the localization quality. Localization accuracy and repeated localization accuracy are measured and shown in box plot in Fig. 7b. The accuracy and robustness of localization have been improved, as evidenced by lower mean and interquartile range (IQR) values.

IV. Conclusion And Future Works

This paper presents PS-Loc, a robust optimal transport-based framework for prior structural reference-based LiDAR localization. This framework features both OT matching, a novel matching approach that considers surrounding structural relations, and PG-PA, a structural reference updating strategy that combines prior and observation information. These features improve the localization precision and robustness to rapid rotation and dynamics. The results demonstrate that this framework can improve the localization precision in rapid rotation and dynamic scenes, and even outperforms LiDAR-Inertial based odometry in a certain dynamic environments. The framework can be integrated into other SLAM system, enhancing localization quality without the need for high-precision maps. In future works, the robustness to degraded environments with all parallel walls need considering, avoiding inaccurate pose estimation through adaptive structural feature weighting.

References

[1] J. Zhang and S. Singh, "LOAM: Lidar odometry and mapping in real-time." in *Robotics: Science and systems*, vol. 2, no. 9. Berkeley, CA, 2014, pp. 1–9.

[2] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized LiDAR odometry and mapping on variable terrain," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4758–4765.

[3] 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.

[4] F. Boniardi, T. Caselitz, R. Kümmerle, and W. Burgard, "A pose graph-based localization system for long-term navigation in CAD floor plans," *Robotics and Autonomous Systems*, vol. 112, pp. 84–97, 2019.

[5] M. Oelsch, M. Karimi, and E. Steinbach, "RO-LOAM: 3d reference object-based trajectory and map optimization in LiDAR odometry and mapping," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 6806–6813, 2022.

[6] Y. Wang, Y. Fan, J. Wang, and W. Chen, "Long-term navigation for autonomous robots based on spatio-temporal map prediction," *Robotics and Autonomous Systems*, vol. 179, p. 104724, 2024.

[7] F. Boniardi, T. Caselitz, R. Kümmerle, and W. Burgard, "Robust LiDAR-based localization in architectural floor plans," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 3318–3324.

[8] M. Mielle, M. Magnusson, and A. J. Lilienthal, "The auto-complete graph: Merging and mutual correction of sensor and prior maps for SLAM," *Robotics*, vol. 8, no. 2, p. 40, 2019.

[9] M. Oelsch, M. Karimi, and E. Steinbach, "R-loam: Improving lidar odometry and mapping with point-to-mesh features of a known 3d reference object," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2068–2075, 2021.

[10] D. Xu, J. Liu, J. Hyyppä, Y. Liang, and W. Tao, "A heterogeneous 3D map-based place recognition solution using virtual LiDAR and a polar grid height coding image descriptor," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 183, pp. 1–18, 2022.

[11] L. Gao and L. Kneip, "FP-Loc: Lightweight and drift-free floor plan-assisted LiDAR localization," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 4142–4148.

[12] J. Luo, Q. Ye, S. Zhang, and Z. Yang, "Indoor mapping using low-cost MLS point clouds and architectural skeleton constraints," *Automation in Construction*, vol. 150, p. 104837, 2023.

[13] G. Ferrer, "Eigen-factors: Plane estimation for multi-frame and time-continuous point cloud alignment," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 1278–1284.

[14] L. Zhou, D. Koppel, and M. Kaess, "LiDAR SLAM with plane adjustment for indoor environment," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7073–7080, 2021.

[15] M. Xu, S. Lin, J. Wang, and Z. Chen, "A LiDAR SLAM system with geometry feature group based stable feature selection and three-stage loop closure optimization," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[16] Z. Liu and F. Zhang, "Balm: Bundle adjustment for LiDAR mapping," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3184–3191, 2021.

[17] Z. Liu, X. Liu, and F. Zhang, "Efficient and consistent bundle adjustment on lidar point clouds," *IEEE Transactions on Robotics*, 2023.

[18] X. Du, Y. Lu, and Q. Chen, "A fast multiplane segmentation algorithm for sparse 3-d LiDAR point clouds by line segment grouping," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–15, 2023.

[19] L. V. Kantorovich and S. Rubinshtein, "On a space of totally additive functions," *Vestnik of the St. Petersburg University: Mathematics*, vol. 13, no. 7, pp. 52–59, 1958.

[20] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," in *Sensor fusion IV: control paradigms and data structures*, vol. 1611. Spie, 1992, pp. 586–606.

[21] Y. Gao and M. K. Leung, "Line segment Hausdorff distance on face matching," *Pattern Recognition*, vol. 35, no. 2, pp. 361–371, 2002.

[22] R. Sinkhorn, "A relationship between arbitrary positive matrices and doubly stochastic matrices," *The annals of mathematical statistics*, vol. 35, no. 2, pp. 876–879, 1964.

[23] T. Deng, Y. Wang, H. Xie, H. Wang, J. Wang, D. Wang, and W. Chen, "Neslam: Neural implicit mapping and self-supervised feature tracking with depth completion and denoising," *arXiv preprint arXiv:2403.20034*, 2024.

[24] B. Peng, H. Xie, and W. Chen, "Roll: Long-term robust LiDAR-based localization with temporary mapping in changing environments," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 2841–2847.

[25] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "Fast-lio2: Fast direct LiDAR-inertial odometry," *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2053–2073, 2022.