Proceedings of the 2021 IEEE
International Conference on Robotics and Biomimetics
December 27-31, 2021, Sanya, China

2021 IEEE International Conference on Robotics and Biomimetics (ROBIO) | 978-1-6654-0535-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ROBIO54168.2021.9739599

Pole-like Objects Mapping and Long-Term Robot Localization in Dynamic Urban Scenarios

Zhihao Wang, Silin Li, Ming Cao, Haoyao Chen*, and Yunhui Liu

***Abstract*— Localization on 3D data is a challenging task for unmanned vehicles, especially in long-term dynamic urban scenarios. Due to the generality and long-term stability, the pole-like objects are very suitable as landmarks for unmanned vehicle localization in time-varying scenarios. In this paper, a long-term LiDAR-only localization algorithm based on semantic cluster map is proposed. At first, the Convolutional Neural Network(CNN) is used to infer the semantics of LiDAR point clouds. Combined with the point cloud segmentation, the static objects pole/trunk are extracted and registered into global semantic cluster map. When the unmanned vehicle re-enters the environment again, the relocalization is completed by matching the clusters of current scan with the clusters of the global map. Furthermore, the matching between the local and global maps stably outputs the global pose at 2Hz to correct the drift of the 3D LiDAR odometry. The experimental results on our campus dataset demonstrate that the proposed approach performs better in localization accuracy compared with the current state-of-the-art methods. The source of this paper is available at: http://www.github.com/HITSZ-NRSL/long-term-localization.**

I. Introduction

Simultaneous Localization and Mapping (SLAM) can provide location and map information for unmanned vehicles in unknown environments. Many classical SLAM methods [1], [2] have been proposed and applied in static environments. However, the real urban environments are time-varying, such as, lighting changing, weathers and seasons varying, and the movement of objects (e.g., cars parked on both sides of the road may be driven away after a while), etc. And most descriptors do not take the large environmental changes into account. Thus, the dynamic environments introduce difficulties for an unmanned vehicle to localize itself when it re-enters a known and mapped environment. Some literatures [3] rely on regularly updating high-definition point cloud maps to deal with the problem of dynamic environments, which is time and computation consuming. Pole-like objects, such as street lamps, poles of building and tree trunks, etc., are ubiquitous in urban areas [4]; they are long-term stable and invariant, which make pole-like objects suitable as landmarks for accurate and reliable relocalization. But this work ignored the geometric relationships between pole-like objects in localization.

Compared with the low-level point features, it is a more robust way to describe the environmental changes by using the high-level semantic features. The structure of the environment is stable, so many solutions based on laser sensors have been researched. Scan Context Image(SCI) [5] is proposed to encode the top view of LiDAR point clouds into a three-channel color image, then the place recognition problem is converted to a classification problem. A novel semantic graph [6] is proposed for point clouds by reserving the semantic and topological information of the raw point clouds. Thus, place recognition is modeled as graph matching problem. However, this work has not been verified on the long-term place recognition problem. A global localization algorithm [7] is proposed by using only a single 3D LiDAR scan at a time, which relies on learning-based descriptors of point cloud segments. But its performance on long-term dataset is underperformance and not stable.

Thus, an approach based on semantic cluster is presented for long-term relocalization in urban dynamic environments, relying on pole/trunk landmarks extracted from mobile LiDAR data. To summarize, the main contributions of this paper are three-fold.

- A method to extract semantic cluster of pole-like

This study was supported in part by the National Natural Science Foundation of China under Grants U1713206, the Shenzhen Science and Innovation Committee under Grants JCYJ20200109113412326, JCYJ20180507183837726, and JCYJ20180507183456108, and the National Key Research and Development Program of China under Grant 2018YFB1309300.

Z.H. Wang, S.L. Li, M. Cao, and H.Y. Chen(*Corresponding author) are with the School of Mechanical Engineering and Automation, Harbin Institute of Technology Shenzhen, P.R. China, e-mail: hychen5@hit.edu.cn.

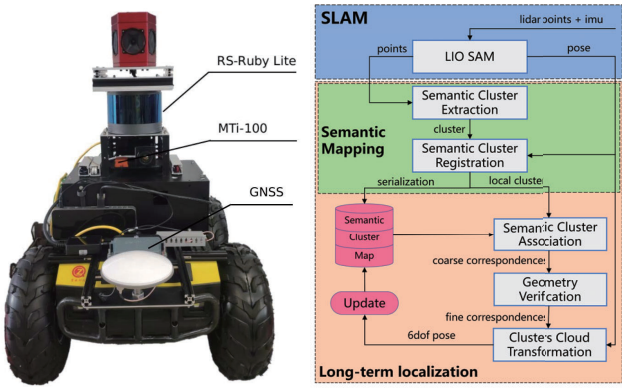
Y.H. Liu is with the Department of Mechanical and Automation Engineering, Chinese University of Hong Kong, P.R. China, e-mail: yhliu@mae.cuhk.edu.hk.

objects from raw 3D LiDAR points and create a robust semantic cluster map is proposed to resolve the long-term challenge.

- A semantic cluster association algorithm based on geometric consistency is proposed to relocalize unmanned vehicles in long-term scenarios.
- A long-term and real-time localization system is developed based on the robust semantic cluster relocalization module.

II. System Overview

The experimental unmanned vehicle system is shown in Fig. 1(a). The RS-Ruby-Lite 3D LiDAR with 80 laser-beam has a vertical field of view (FOV) of 40° and a horizontal FOV of 360°. And the inertial sensor MTi-100 is also equipped with the unmanned vehicle MR1000. Global Navigation Satellite System (GNSS)/Inertial Navigation System(INS) are used as the ground truth for comparison.



(a) Research platform (b) System overview

Fig. 1. Platform and system overview of long-term localization algorithm

The framework of the long-term localization system is shown as Fig. 1(b). The semantic clusters of pole-like objects are extracted by the semantic labels of the point clouds, which are obtained directly through CNN. Then, the semantic clusters are registered in the global semantic cluster map using the six-degree-of-freedom(6-DOF) pose obtained by SLAM module. When the unmanned vehicle re-enters the environment, the local semantic cluster map is built by current scan. Then, the association algorithm based on the semantic labels and the geometric structure is proposed to calculate the pose transformation between the local and the global map. And the relocalization module continuously runs and stably outputs the global pose at 2Hz to correct the drift of the LiDAR odometry.

III. Method

A. Semantic Cluster Extraction

The semantic labels of LiDAR point clouds are annotated on self-made campus dataset and the network of RangeNet++ [8] is retrained for inferring the semantic labels of point clouds. The range-image based method [9] is used to segment the point clouds into clusters based on the semantic labels. For each cluster, the label of cluster is voted by the statistical number of the point labels in the cluster, by considering the imperfection of point cloud semantic segmentation.

B. Semantic Cluster Registration

The registration means inserting semantic clusters into a map, and the map can be global or local semantic cluster map. The global map represents the geometric information of pole-like objects in history environments, such as the environment of several months ago. The local map represents the geometric information of current scan. This section only describes the registration of the global cluster map in details, while the registration of local map is similar.

Let $\mathcal{C}$ denotes the semantic cluster, and $\mathcal{M}_g^{cur}$ denotes the global cluster map built on scan frames from t_0 to t_{k-1}. $\mathcal{M}_g^{cur}$ contains j clusters, denoted as

$$\mathcal{M}_g^{cur} = \{\mathcal{C}_{g1}, \mathcal{C}_{g2}, ..., \mathcal{C}_{gj}\} \tag{1}$$

For the building process, the global map is initialized by adding all the semantic clusters of the frame at t_0. After initialization, the clusters of current frame at t_k need to be added. Specifically, for each cluster $\mathcal{C}_{gk}$ in current frame at t_k, the position is transformed into the coordinate system of the global map by using LIO SAM [10]. The label of $\mathcal{C}_{gk}$ is set as the same with the searched closest neighbor cluster using k nearest neighbor(KNN) [11]. And if the closest neighbor cluster of $\mathcal{C}_{gk}$ is not found, $\mathcal{C}_{gk}$ is directly inserted into the global map as a new cluster $\mathcal{C}_{g\{j+1\}}$.

The registered 3D semantic cluster map is shown in Fig. 2(a). The clusters mainly include poles and trunks, and they are perpendicular to the horizontal plane in the environment. Therefore, the 2D centroid points are able to represent the geometric information of the clusters, and are registered into the cluster map. The registered 2D semantic cluster centroids are the XY-plane projection of 3D centroids, as shown in Fig. 2(b).

C. Semantic Cluster Association

The geometric information of the cluster to be matched is described as the relative positions and

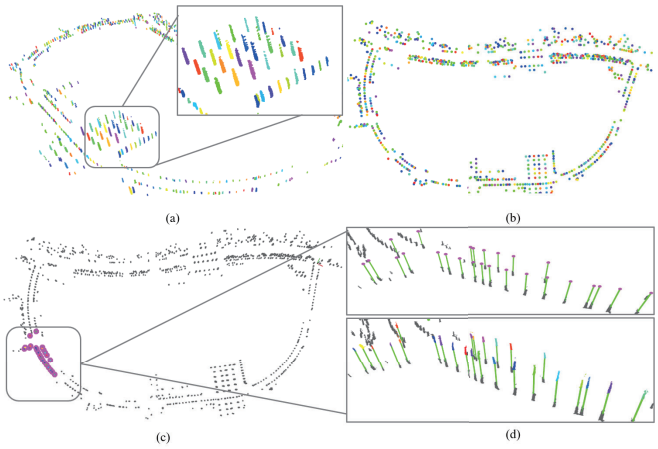


Fig. 2. An illustration of the long-term relocalization approach. The 3D semantic cluster map is shown in (a), each color represents a semantic cluster. The 2D semantic cluster map is the top view of 3D cluster map in (b), and each circle represents 3D centroid point of a semantic cluster. In (c), gray dots represent the clusters of the global semantic cluster map, pink dots represent the clusters of the local semantic cluster map. In (d), the green lines show the matched pairs between the local and global map. The upper figure shows the matched pairs of the 2D centroid points, while the lower figure shows the matched pairs of the 3D centroid points.

angles with its neighbors. Let $\mathcal{M}_l$ denotes the local semantic cluster map with P clusters extracted from one scan. The registered global semantic cluster map $\mathcal{M}_g$ contains Q clusters. And the definitions are shown as follows:

$$\begin{aligned}\mathcal{M}_l &= \{\mathcal{C}_{l1}, \mathcal{C}_{l2}, ..., \mathcal{C}_{li}, ..., \mathcal{C}_{lP}\} \\ \mathcal{M}_g &= \{\mathcal{C}_{g1}, \mathcal{C}_{g2}, ..., \mathcal{C}_{gj}, ..., \mathcal{C}_{gQ}\}\end{aligned} \tag{2}$$

In the local map $\mathcal{M}_l$, the neighbor clusters of $\mathcal{C}_{li}$ within the search radius are denoted as $\{\mathcal{N}_{\mathcal{C}_{li}} = \mathcal{C}^i_{li}, i = 1, 2, ..., k_l\}$. Similarly, $\{\mathcal{N}_{\mathcal{C}_{gj}} = \mathcal{C}^j_{gj}, j = 1, 2, ..., k_g\}$ denotes the neighbor clusters of $\mathcal{C}_{gj}$ within the search radius in $\mathcal{M}_g$. And the number of neighbor clusters is k_l and k_g, respectively.

The edges are the lines between $\mathcal{C}_{li}$ and its neighbor clusters in $\mathcal{N}_{\mathcal{C}_{li}}$, and these edges form the edge set $\{E_{\mathcal{C}_{li}} = e^i_l, i = 1, 2, ..., k_l\}$. Similarly, $\{E_{\mathcal{C}_{gj}} = e^j_g, j = 1, 2, ..., k_g\}$ denotes the cluster edge set between $\mathcal{C}_{gj}$ and its neighbor clusters in $\mathcal{N}_{\mathcal{C}_{gj}}$. The principles for matching two clusters are the same labels and enough matched edges. Then, the matching problem of semantic clusters is transformed into the matching of edges in the cluster edge sets.

1) Edge Association: For the matching process of e^i_l and e^j_g, the length of each edge in $E_{\mathcal{C}_{gj}}$ is calculated, and the top n edges whose length are nearest to the length of e^i_l are selected as candidate edges of e^i_l, denoted as $\{{}^i\hat{e}^1_g, {}^i\hat{e}^2_g, ..., {}^i\hat{e}^j_g, ..., {}^i\hat{e}^n_g\}$, and n is a constant(default: 5). $(e^i_l, {}^i\hat{e}^j_g)$ denotes the candidate edge pair, and the distance of the candidate edge pair is $\mathcal{D}^{gj}_{li}$, which will be defined later. Thus, the problem of edge matching is formulated as follows:

$$\text{Min}\left(\mathcal{D}^{g1}_{li}, \mathcal{D}^{g2}_{li}, ..., \mathcal{D}^{gj}_{li}, ..., \mathcal{D}^{gn}_{li}\right) < \delta_e \tag{3}$$

where δ_e is the maximum tolerance distance for successful matching of edge pairs.

For matching e^i_l, not only the length and label of edge e^i_l should be the same, but also the geometric relationships between the edge e^i_l and its neighbors should also be consistent. Therefore, sub-edge is introduced for verifying the neighbor geometric relationships. The sub-edges of e^i_l is the edges in edge set $E_{\mathcal{C}_{li}}$ except e^i_l, and denoted SE_{li}. Similarly, the edges in $E_{\mathcal{C}_{gj}}$ except e^j_g are called sub-edge set SE_{gj} of e^j_g:

$$\begin{aligned}SE_{li} &= \{e^1_l, ..., e^p_l, ..., e^{k_l}_l\}, p \neq i \\ SE_{gj} &= \{e^1_g, ..., e^q_g, ..., e^{k_g}_g\}, q \neq j\end{aligned} \tag{4}$$

The distance $\mathcal{D}^{gj}_{li}$ of the candidate edge pair (e^i_l, e^j_g) is defined as the average distances of successfully matched sub-edge pairs between the sub-edge sets SE_{li} and SE_{gj}.

2) Sub-Edge Association: The sub-edge e^p_l in SE_{li} includes elements $\left[d^p_l, \theta^p_l\right]$, where d^p_l is the length of e^p_l, and θ^p_l is the clockwise angle between e^p_l and e^i_l. As the same, $[d^q_g, \theta^q_g]$ denotes the elements of sub-edge e^q_g in SE_{gj}. A cluster association example is shown in Fig. 3.

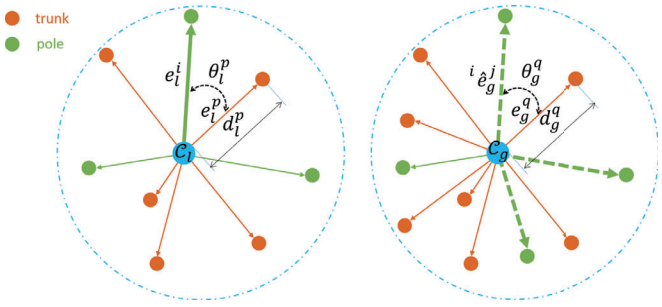


Fig. 3. Cluster association diagram. (a) shows the neighbor semantic clusters of $\mathcal{C}_{li}$ in local semantic cluster map, and (b) shows the neighbors of $\mathcal{C}_{gj}$ in the global cluster map. Thick solid line with arrows in (a) represents the edge to be matched, and thin solid lines with arrows represent the sub-edges. θ, d represent the element of sub-edge. The thick dotted lines in (b) indicate the candidate edges of the edge to be matched in (a).

The matching success of two sub-edges e^p_l and e^q_g should satisfy four conditions: (a) the semantic cluster

labels of two sub-edges are the same; (b) The length error between two sub-edges is under a certain threshold; (c) The angle error between two sub-edges is under a certain threshold; (d) The distance between two sub-edges, denoted as d_{lp}^{gq}, is also under a certain threshold. And the distance [12] is defined as follows:

$$d_{lp}^{gq} = \sqrt{d_l^p \cdot d_l^p + d_g^q \cdot d_g^q - 2 \cdot d_l^p \cdot d_g^q \cdot cos(|\theta_l^p - \theta_g^q|)} \tag{5}$$

The four conditions of sub-edges matching are mathematically formulated as follows:

$$\begin{cases} Label(e_l^p) == Label(e_g^q) \\ |d_l^p - d_g^q| < \delta_d \\ |\theta_l^p - \theta_g^q| < \delta_\theta \\ d_{lp}^{gq} < \delta_{se} \end{cases} \tag{6}$$

where δ_* indicates the threshold of sub-edges;

3) Semantic Cluster Association: If more than N_{se} sub-edges of the candidate edge pair (e_l^i, e_g^j) are satisfied with (6), the candidate edge pair is matched successfully. Then, the distance $\mathcal{D}_{li}^{gj}$ of candidate edge pair can be computed as the first line of (7).

$$\mathcal{D}_{li}^{gj} = \begin{cases} log\frac{k_l - 1}{k_{se}^{ij}} \cdot \frac{1}{k_{se}^{ij}} \sum_{pq}^{k_{se}^{ij}} d_{lp}^{gq}, & k_{se}^{ij} >= N_{se} \\ DOUBLE\_MAX, & \text{otherwise} \end{cases} \tag{7}$$

where k_{se}^{ij} indicates the number of sub-edge pairs successfully matched, $k_l - 1$ indicates the number of all sub-edges in SE_{li}, and N_{se} is a threshold. As shown in (7), the distance $\mathcal{D}_{li}^{gj}$ considered the distances of all the matched sub-edges and the proportion of the matched sub-edges k_{se}^{ij} to all sub-edges in SE_{li}. Considering these two factors is to avoid the false matching caused by small average distance of sub-edges and the low proportion of successful matched sub-edges pairs.

From (3) and (7), the best candidate edge pair is formulated as follows:

$$\hat{\mathcal{D}_{li}} = \text{Min}\left(\mathcal{D}_{li}^{g1}, \mathcal{D}_{li}^{g2}, ..., \mathcal{D}_{li}^{gj}, ..., \mathcal{D}_{li}^{gn}\right) < \delta_e \tag{8}$$

where $\hat{\mathcal{D}_{li}}$ is the distance of the best candidate edge pair, δ_e denotes the threshold of the distance.

After matched the candidate edge pairs, the matching of clusters are denoted as follows:

$$\begin{aligned} Label(\mathcal{C}_{li}) == Label(\mathcal{C}_{gj}) \\ k_e >= N_e \end{aligned} \tag{9}$$

where k_e is the numbers of matched edge pairs, and N_e represents the minimum numbers of edge pairs that successfully matched. A matching result example of semantic cluster association is shown in Fig. 2(c) and (d).

D. Long-term Relocalization

Let $c_i = (\mathcal{C}_{li}, \mathcal{C}_{gi})$ denote a matched pair of semantic clusters, where $\mathcal{C}_{li} \in \mathcal{M}_l$ and $\mathcal{C}_{gi} \in \mathcal{M}_g$. Semantic cluster matching pairs obtained from semantic cluster association algorithm are coarse correspondences. Thus, a geometric consistency method [13] will be used to eliminate the false positive matched pairs, and finally fine correspondences are remained.

Then, the 6-DOF pose transformation between the local map $\mathcal{M}_l$ and the global map $\mathcal{M}_g$ is computed by matching the corresponding semantic cluster centroid points and semantic cluster point clouds. Firstly, the 3D centroid points pairs are extracted from semantic cluster matched pairs, and the RANSAC [14] algorithm is used to filter out wrong matched pairs. Secondly, the coarse coordinate transformation matrix $\mathbf{T}_{gl}$ from the local map to the global map is calculated by using iterative closest point (ICP) algorithm.

Thirdly, the local and global point clouds of the matched cluster pairs are respectively extracted from cluster manager. Using the coarse pose transformation $\mathbf{T}_{gl}$ as the initial value, the fine transformation matrix $\widehat{\mathbf{T}}_{gl}$ is obtained by the ICP algorithm again with the extracted point clouds instead of the 3D centroid points, shown as follows:

$$\widehat{\mathbf{T}}_{gl} = \arg\min_{\widehat{\mathbf{T}}_{gl}} \frac{1}{2} \sum^{k} \left\| P_{li} - \widehat{\mathbf{T}}_{gl} P_{gj} \right\|^2 \tag{10}$$

where P_{li}, P_{gj} are the extracted point clouds of the clusters $\mathcal{C}_{li}$ and $\mathcal{C}_{gj}$; $\widehat{\mathbf{T}}_{gl}$ represents the global pose of the local map in the global map.

E. Long-term Localization

The global pose $\widehat{\mathbf{T}}_{gl}$, stably output at 2Hz by the relocalization module, is used to correct the accumulated drift of the LiDAR odometry in real time. The global pose $\widehat{\mathbf{T}}_{gl}$ at t_k is re-expressed as T_k^{WL}. The LiDAR odometry outputs n poses continuously from t_k to t_{k+n} before the next global pose is updated, denoted as $\{T_k^L, T_{k+1}^L, ..., T_{k+i}^L, ..., T_{k+n-1}^L\}$. During t_k to t_{k+n}, the global pose estimation of the unmanned vehicle is calculated as follows:

$$T_{k+i}^{WL} = T_k^{WL} \cdot \prod_{j=0}^{i} T_{k+j}^L \tag{11}$$

where $i \in [0, n-1]$. So far, long-term localization has been finished by correcting the drift of LiDAR odometry with the real-time global pose of relocalization module.

IV. Experiments

To evaluate the proposed relocalization and localization algorithms in long-term scenarios, several experiments were performed on the self-made campus dataset. It is worth noted that Carlevaris-Bianco [15] proposed NCLT dataset in long-term scene, but the Velodyne32 LiDAR in NCLT dataset is installed upside down, which makes the actual LiDAR point cloud sparse. While the proposed method is based on semantic segmentation of dense LiDAR point clouds, NCLT dataset cannot meet our requirements. Therefore, the proposed method is compared on the self-made dataset with the SCI algorithm [5], which is the state-of-the-art long-term LiDAR-based localization system.

A. Benchmark Dataset

The challenge of long-term localization is the environmental changes. So the self-made dataset recorded the data in the university town environment with high traffic volume and dense vegetation.

To use RangeNet++ to infer the semantic label, the self-made dataset is annotated with more than 1500 frames of 3D point clouds generated by LiDAR scanning, which covered a geographic area of 76,000 m^2 and included 120 million three-dimensional semantic points. And the marked labels with 10 categories included buildings, highways, sidewalks, bicycles, green plants, poles, tree trunks, vehicles, pedestrians, and other objects. The organization of the semantic point clouds is labeled and organized according to the SemanticKitti [16]. Four datasets were collected, and Table I showed the details of datasets. Although the datasets were collected within a month, some places of the environment have changed significantly due to the movement of cars and growing of trees; furthermore, different numbers of clusters in the global map were randomly removed to simulate the significant changes of environment. We will also continue collect more datasets and update the results in the future.

TABLE I

List of campus data sets.

Date	Length(km)	Time	Weather
2020-10-12	1.31	Afternoon	Sunny
2020-10-18	1.28	Morning	Cloudy
2020-10-25	1.53	Evening	Cloudy
2020-11-05	1.4	Afternoon	Sunny

B. Evaluation of Long-term Relocalization

To evaluate the success rate of relocalization in long-term scenarios, the earliest dataset 2020-10-12 was used to build the global semantic cluster map, called source dataset. The other three datasets were used to create the local semantic cluster maps for relocalization, called match dataset. The definition of successful relocalization is the distance between the estimated position of relocalization and the ground truth by GNSS smaller than the tolerance error $\delta_{\mathrm{r}} = 10m$.

The success rate of relocalization of the SCI algorithm are divided into two categories, namely the success rate of the highest(top 1) scoring frame and the top 5 scoring frames. 120 positions were selected for the relocalization experiments on each dataset. And the success rate of relocalization was shown in Table II. The left number of each cell represents the number of successful relocalizations in 120 experiments, and the right number of each cell represents the success rate. The results showed the success rate of the proposed relocalization algorithm higher than the SCI algorithm on the self-made dataset.

TABLE II

Success rate of SCI and the proposed method.

Dataset	SCI(TOP1)	SCI(TOP5)	Proposed
2020-10-18	99/82.5%	114/95.0%	**120/100.0**%
2020-10-25	97/80.8%	108/90.0%	**116/96.6**%
2020-11-05	103/85.8%	112/93.3%	**118/98.3**%

C. Long-term Localization Analysis

The evaluation of long-term localization experiment is mainly to compare the position estimation accuracy of the LiDAR odometry. The metrics of Root Mean Square Error(RMSE) is used for evaluation, and the results of the proposed long-term localization system and LOAM are compared with the ground truth. The results were shown in Table III. The proposed approach does not utilize the IMU data in localization, therefore it was compared with the LOAM but not LIO-SAM.

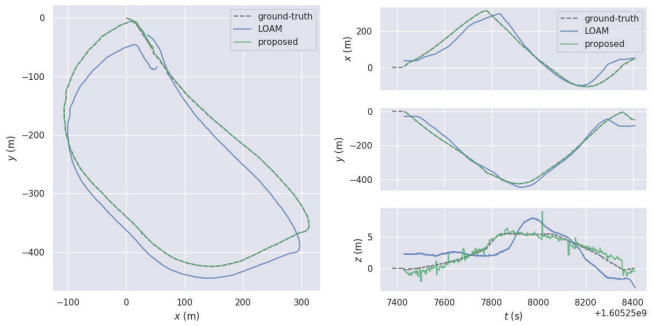
Since LOAM only relies on the LiDAR odometry, the RMSE was relatively large because of the accumulated error of the odometry. For the proposed method, the relocalization algorithm based on the semantic cluster map was used to correcting the accumulated error, and the position estimation between frames of the

TABLE III

RMSE OF LONG-TERM LOCALIZATION.

Dataset	LOAM(m)	Proposed(m)
2020-10-18	42.41	**0.29**
2020-10-25	44.32	**0.34**
2020-11-05	40.48	**0.24**

laser odometry is accurate, therefore the RMSE of the proposed long-term localization system is smaller. The trajectories of ground truth, LOAM and the proposed method are shown in Fig. 4(a) by using the dataset 2020-11-05 as an example, and the x, y, z position error are shown in Fig. 4(b).



(a) (b)

Fig. 4. (a) Trajectories of the ground truth, LOAM and the proposed long-term localization system. (b) Position errors along x-y-z axes within a time window.

V. CONCLUSION

A novel relocalization method based on semantic cluster map is proposed to realize high accuracy relocalization and real-time localization in long-term environments. A robust semantic cluster map built by extracting pole-like objects can resolve the long-term challenge. The association algorithm on the basis of geometric consistency verification for localization showed high accuracy. And a long-term and real-time localization system is developed based on the proposed relocalization module and LiDAR odometry. Finally, several experiments are performed to illustrate the effectiveness of the proposed approach. In the future work, we will extend to utilize more objects in long-term environments and make it more general rather than relying heavily on pole-like objects.

REFERENCES

[1] Ji Zhang and Sanjiv Singh. Loam: Lidar odometry and mapping in real-time. In *Robotics: Science and Systems*, volume 2, 2014.

[2] Tixiao Shan and Brendan Englot. Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4758–4765. IEEE, 2018.

[3] François Pomerleau, Philipp Krüsi, Francis Colas, Paul Furgale, and Roland Siegwart. Long-term 3d map maintenance in dynamic environments. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3712–3719. IEEE, 2014.

[4] Alexander Schaefer, Daniel Büscher, Johan Vertens, Lukas Luft, and Wolfram Burgard. Long-term vehicle localization in urban environments based on pole landmarks extracted from 3-d lidar scans. *Robotics and Autonomous Systems*, 136:103709, 2021.

[5] Giseop Kim, Byungjae Park, and Ayoung Kim. 1-Day Learning, 1-Year Localization: Long-Term LiDAR Localization Using Scan Context Image. *IEEE Robotics and Automation Letters*, 4(2):1948–1955, 2019.

[6] Xin Kong, Xuemeng Yang, Guangyao Zhai, Xiangrui Zhao, Xianfang Zeng, Mengmeng Wang, Yong Liu, Wanlong Li, and Feng Wen. Semantic graph based place recognition for point clouds. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8216–8223. IEEE, 2020.

[7] S. Ratz, M. Dymczyk, R. Siegwart, and R. Dubé. Oneshot global localization: Instant lidar-visual pose estimation. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5415–5421, 2020.

[8] A. Milioto, I. Vizzo, J. Behley, and C. Stachniss. RangeNet++: Fast and Accurate LiDAR Semantic Segmentation. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2019.

[9] I. Bogoslavskyi and C. Stachniss. Fast range image-based segmentation of sparse 3d laser scans for online operation. In *Proc. of The International Conference on Intelligent Robots and Systems (IROS)*, 2016.

[10] Tixiao Shan, Brendan Englot, Drew Meyers, Wei Wang, Carlo Ratti, and Rus Daniela. Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5135–5142. IEEE, 2020.

[11] R.B. Rusu and S. Cousins. 3d is here: Point cloud library (pcl). In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 1 –4, May 2011.

[12] Zhichen Pan, Haoyao Chen, Silin Li, and Yunhui Liu. Clustermap building and relocalization in urban environments for unmanned vehicles. *Sensors*, 19(19):4252, 2019.

[13] Mattia G Gollub, Renaud Dubé, Hannes Sommer, Igor Gilitschenski, and Roland Siegwart. A partitioned approach for efficient graph–based place recognition. In *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst./Workshop Planning, Perception Navigat. Intell. Veh.*, 2017.

[14] Konstantinos G Derpanis. Overview of the ransac algorithm. *Image Rochester NY*, 4(1):2–3, 2010.

[15] Nicholas Carlevaris-Bianco, Arash K Ushani, and Ryan M Eustice. University of michigan north campus long-term vision and lidar dataset. *The International Journal of Robotics Research*, 35(9):1023–1035, 2016.

[16] Jens Behley, Martin Garbade, Andres Milioto, Jan Quenzel, Sven Behnke, Cyrill Stachniss, and Jurgen Gall. Semantickitti: A dataset for semantic scene understanding of lidar sequences. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9297–9307, 2019.