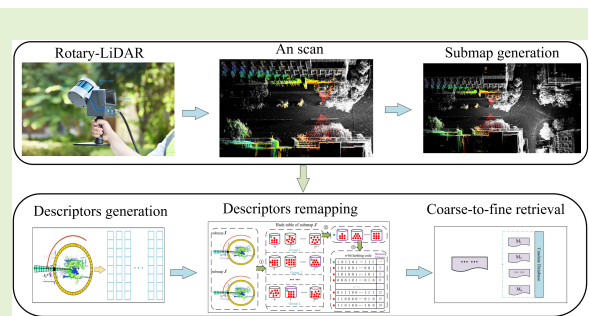


RLS-LCD: An Efficient Loop Closure Detection for Rotary-LiDAR Scans

Qiyuan Zhang, Shunyi Zheng, Rui Li, Xiqi Wang, Yuan He, Xiaonan Wang

Abstract—Loop closure detection is a crucial technology within the field of simultaneous localization and mapping (SLAM), aiming to mitigate cumulative errors and maintain a globally consistent map. Traditional fixed-LiDAR scans have demonstrated reliable loop closure detection (LCD) performance when employing global descriptors. However, for rotary-LiDAR scans, these methods fall short in place recognition due to significant view-angle changes. To address this challenge, we propose a novel LCD algorithm tailored for rotary-LiDAR scans, named RLS-LCD. Our RLS-LCD approach utilizes submap-based recognition and a novel lightweight global descriptor to mitigate the impact of large view-angle changes inherent to rotary-LiDAR scans. Additionally, we adopt a coarse-to-fine recognition strategy to enhance recall and precision of place recognition in similar structure scenes. The experimental findings unequivocally demonstrate that the RLS-LCD algorithm not only outperforms the contemporary LCD methodologies reliant on global descriptors but also exhibits the capability to execute in real-time on computing platforms with limited computational resources.

Index Terms—Loop Closure Detection, Intensity Coding, Coarse-to-fine Recognition, Rotary-LiDAR, LiDAR SLAM.



The experimental findings unequivocally demonstrate that the RLS-LCD algorithm not only outperforms the contemporary LCD methodologies reliant on global descriptors but also exhibits the capability to execute in real-time on computing platforms with limited computational resources.

I. INTRODUCTION

DRIVEN by the advances in sensor technologies and the continuous maturity of relevant optimization theories, simultaneous localization and mapping (SLAM) has experienced rapid development in recent years. Benefiting from the low deployment requirements and the advantage of real-time positioning, the SLAM technology plays a crux role in many practical applications such as autonomous driving [1], robot navigation and mobile mapping [2]. For the integrated 3D LiDAR mobile mapping based on SLAM, the whole system needs to estimate ego-motion in an unknown environment and construct the map at the same time. Therefore, the pose estimation error is prone to accumulate over time, resulting in trajectory drift and structural error after long-distance travel [3]. An effective way to eliminate the cumulative error is to perform a loop closure optimization after running for a period of time [4], which can not only improve positioning accuracy but also help to establish a globally consistent map. The first step in loop closure optimization is loop closure detection (LCD), which is to identify and match the visited place [5].

(Corresponding author: Shunyi Zheng)

Qiyuan Zhang, Shunyi Zheng, and Xiqi Wang are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, 430079, Hubei Province, China (e-mail: qiyuanzhang@whu.edu.cn; syzheng@whu.edu.cn; wangxiqi@whu.edu.cn).

Rui Li is with Intelligent Control & Smart Energy (ICSE) Research Group, School of Engineering, University of Warwick, Coventry, CV4 7A1, UK (e-mail: rui.li.4@warwick.ac.uk).

Yuan He, and Xiaonan Wang are with Wuhan ZG Automation Technology Co., Ltd., Wuhan, 430079, Hubei Province, China (e-mail: yuan.he@hexagon.com; xiaonan.x.wang@hexagon.com).

In recent years, a variety of effective LCD methods have been proposed for LiDAR SLAM and integrated 3D LiDAR mobile mapping. Basically, the methodology of these methods is to generate descriptors for each LiDAR scan and then to identify the loop closure candidate from the historical scans by measuring the similarity between the descriptors of the query and the candidate scans. According to the similarity measurement techniques, the existing place recognition methods can be divided into three categories: local descriptors based [6]–[8], global descriptors based [4], [9]–[11], and learning-based methods [1], [12]–[14]. The LCD methods have been widely employed for fixed-LiDAR scans and have achieved remarkable performance in both efficiency and accuracy [15]–[17]. However, the application of LCD for rotary-LiDAR scans has not been reported yet. In recent years, there have been several successful commercial products for mobile mapping based on rotary-LiDAR on the market (e.g. GeoSLAM [18], RigelsLAM [19], Lixel L1 [20]), but few pieces of literature have studied the technology of mobile mapping based on rotary-LiDAR (such as place recognition based on rotary-LiDAR scans).

Why do we need to exploit rotary-LiDAR for mapping? As shown in Fig. 1(a), the fixed-LiDAR is the 3D LiDAR installed on a fixed head. During the scanning process, the laser transmitter in the LiDAR rotates horizontally only, while the 3D LiDAR itself does not rotate. Therefore, limited by the vertical field-of-view (FoV) of the 3D LiDAR, the integrity of mapping by the fixed-LiDAR scanning is not satisfactory (Fig. 2(a)). Rotary-LiDAR (Fig. 1(b)) refers to the 3D LiDAR installed on a rotary head, and the motor drives 3D LiDAR to

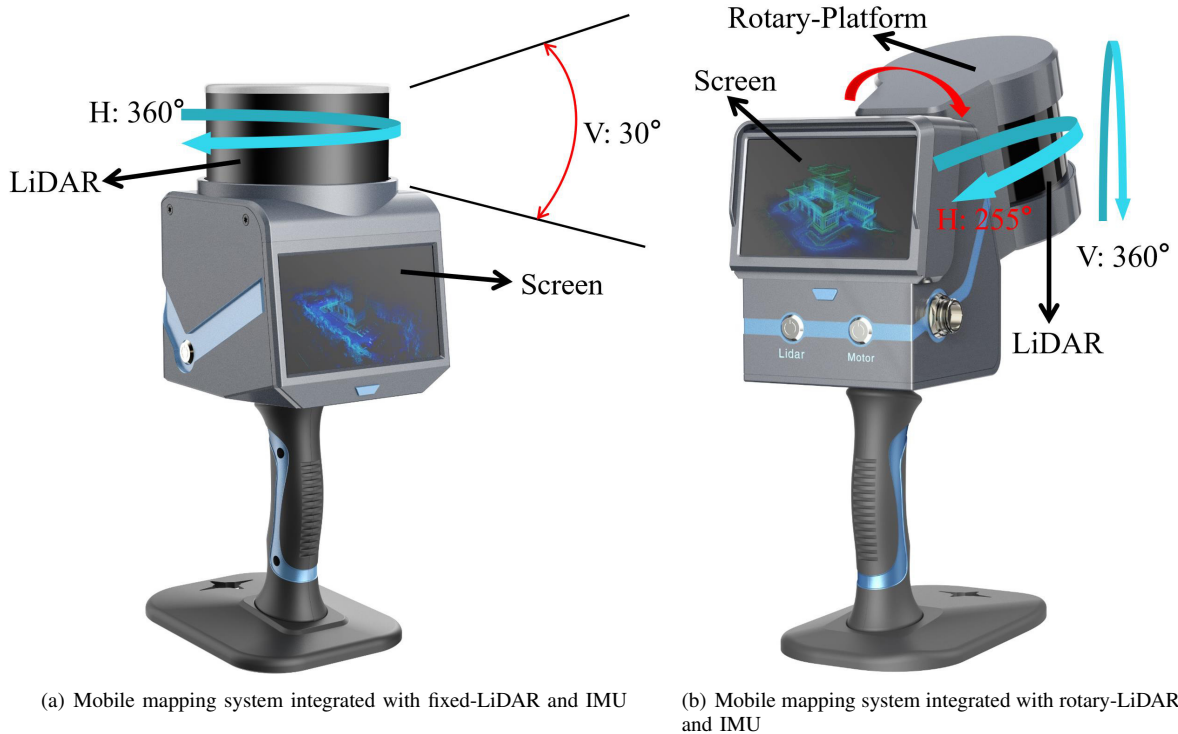


Fig. 1. Two kinds of developed mobile mapping systems. The horizontal FoV of the fixed-LiDAR is 360°, and the vertical FoV is 30°. When the rotary platform does not rotate, the horizontal FoV of the rotary-LiDAR is 255° and the vertical FoV is 30°. When the rotary platform rotates, the vertical FoV of rotary-LiDAR can reach 360°.

rotating around a specific rotation shaft at a fixed frequency, thereby expanding the scanning FoV of the 3D LiDAR. Therefore, with a larger scanning FoV, the rotary-LiDAR has better integrity for scene scanning (Fig. 2(b)) compared with fixed-LiDAR. As far as we know, due to the advantage of complete mapping, mobile mapping technology based on rotary-LiDAR has received a lot of attention in the fields of urban 3D digitization, tunnel 3D mapping, and so on. However, the great change in rotary-LiDAR scanning view-angle also brings new and huge challenges to place recognition. Specifically, the ever-changing view-angles of rotary-LiDAR scans cause the low overlap of scanned scenes when visiting the same location at different times (Fig. 3). Therefore, the performance of the traditional scan-based LCD method for rotary-LiDAR scan place recognition are far from satisfactory.

In this paper, we propose a lightweight LCD algorithm for rotary-LiDAR scans, i.e. RLS-LCD, and then integrate it into the developed mobile mapping system¹ (Fig. 1(b)). Compared to traditional scan-based methods, RLS-LCD adopts submap-based place recognition and it adopts the lightweight and unique descriptor we designed, which can significantly reduce the impact of large changes in scanning view-angle of rotary-LiDAR. RLS-LCD also adopts a coarse-to-fine recognition strategy, which can improve recognition accuracy in single structured scenes.

The main contributions of this paper are as follows:

- 1) We propose a lightweight place recognition method suitable for rotary-LiDAR scans and integrate the algorithm

into the mobile mapping system.

- 2) We utilize submap-based recognition to reduce the adverse impact of large changes in the view-angle of rotary-LiDAR scanning on place recognition. Moreover, we design a novel global descriptor tailored to encode the intricate intensity information.
- 3) We design a coarse-to-fine place recognition strategy to improve the precision of the loop closure detection based on global descriptors for similar scene recognition.

II. RELATED WORK

For the SLAM-based mobile mapping technology, loop closure optimization is a crucial technology in mobile mapping, which can effectively eliminate accumulated errors and enable map building with global consistency [5], [21]. The premise of loop closure optimization is to recognize a previously visited location. To this end, the LCD methods for fixed-LiDAR scans normally calculate the similarity of the descriptors between the query scan and the candidate scan first and then select the matched positions accordingly. According to the principle of generation, descriptors can be divided into local descriptors and global descriptors [15]. In [6], the fast point feature histogram (FPFH), a typical local descriptor, was employed as a histogram to encode the local normal vector and curvature features of point clouds. Bosse et al. [22] proposed a probabilistic voting strategy, which utilizes Gasalt3D descriptors to express local features. In addition to the local geometric features, the intensity information of LiDAR point clouds can also be employed to generate local descriptors. For example, Guo et al. [8] designed a novel local descriptor called intensity

¹<https://youtu.be/x9y5pfK7c5w>

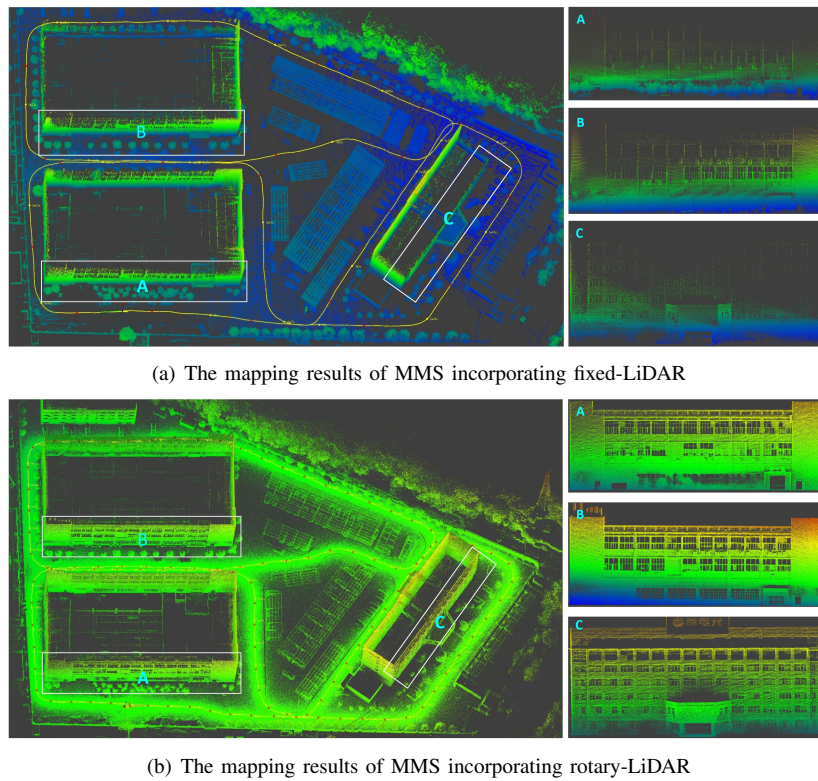


Fig. 2. Comparison of the mapping results of two types of mobile mapping systems for the same scene (the scanning trajectory is indicated by the solid yellow line). When comparing areas *A*, *B* and *C*, it becomes apparent that the mobile mapping based on rotary-LiDAR offers distinct advantages in terms of map completeness and level of detail, despite utilizing the same scanning trajectory. Map completeness and detail are crucial metrics for assessing the performance of mobile mapping systems.

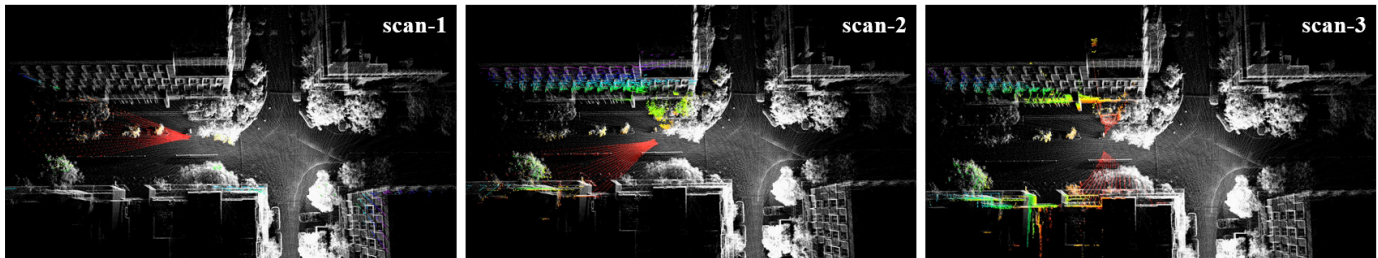


Fig. 3. Rotary-LiDAR scans at different times in the same place. The white point cloud is global map points, the color point cloud is the current scan. We can find that the view-angle of rotary-LiDAR changes greatly and the overlap between scan scenes is low.

signature of histograms of orientations (ISHOT) based on the intensity reading and local geometric features. Shan et al. [23] developed a novel place recognition method based on local descriptors with intensity information. Dubé et al. [24] proposed a novel algorithm called Segmatch, which can generate a special local descriptor based on the segment result of each point cloud generated by the deep learning technology.

Normally, place recognition based on local descriptors needs to extract a large number of local features and calculate descriptors, while local descriptors are easily affected by the change in the view-angle. To address this issue, LCD methods with more stable global descriptors have been developed. In [11] an algorithm called M2DP was proposed which projected 3D point clouds obtained by LiDAR scans onto a 2D plane to extract global descriptors. Kim et al. [4] designed a scan context (SC) algorithm. In their pipeline, the laser scan was first projected to a circular 2D plane according to the polar

coordinates, where the 2D plane was divided into a grid by a certain angle and radial edge length. Secondly, the maximum Z-value of the 3D point in each grid was taken as the value of the global descriptor element. Finally, the grid was expanded to get a global descriptor similar to an image. Wang et al. [9] proposed a variation of the SC algorithm called intensity scan context (ISC), which harnessed the intensity of LiDAR scans to express the structural features of the scanned scene. Jiang et al. [17] proposed a simple, lightweight and efficient topological loop closure detection method based on multi-layer contour matching in the bird's eye view (BEV) frame, called Contour Context (Cont2).

In recent years, deep learning technology has also been widely employed in LCD for fixed-LiDAR scans. In [25], based on PointNet [26] and NetVLAD [12], a neural network named PointNetVLAD for place recognition of large-scale LiDAR scans was developed. Liu et al. [27] developed a

neural network called large-scale place description network (LPD-Net) to generate distinguishing and discriminative global descriptors for 3D point clouds. Inspired by the perspective of humans, Kong et al. [13] designed a method that is robust to occlusion and view-angle changes by identifying semantic targets between scans. Li et al. [14] proposed a semantic scan context (SSC) algorithm that uses scan context descriptors [4] to encode high-level semantic features, which improves the performance of place recognition in complex scenes. Kavisha et al. [28] proposed a novel Locus algorithm that was not only invariant to the arrangement of input features but also improved the robustness and distinguishability of representing scan scenes. To be specific, the Locus algorithm extracted and encoded the topology and temporal information of segmented features in scans and aggregated multi-level features by second-order pooling and nonlinear transformation to generate a fixed-length global descriptor. Compared with the traditional methods, the LCD methods based on deep learning can take advantage of high-level features, which are thus more robust to the changes of scan scene and view-angle. However, these methods require additional training steps and higher computational power, which hugely affect their deployment flexibility [16].

When it comes to rotary-LiDAR scans, the low overlap of scanned scenes caused by the ever-changing view-angles will significantly weaken the effectiveness of scan-based methods. As far as we know, there are no LCD methods specifically designed for rotary-LiDAR scans currently. Therefore, the RLS-LCD proposed in this paper would be the first attempt to develop a rotary-LiDAR-oriented LCD method, which can be integrated with the mobile mapping system to enhance the global consistency of mapping.

III. METHODOLOGY

In this section, we will describe the proposed RLS-LCD algorithm in detail. The first step of RLS-LCD is to generate submaps (Section III-A)). Considering that our designed method needs to be integrated into a low computational power mobile mapping system, we designed a lightweight global descriptor to measure the similarity between submaps. To expedite the candidate retrieval process, we use hashing remapping [29] to remapping 2-dimensional global descriptors to 1-dimensional (Section III-B)). To improve the accuracy of recognition in structurally similar scenes (such as underground), in the candidate retrieval step, we designed a coarse-to-refine recognition strategy (Section III-C)). The flow and key steps of our proposed RLS-LCD algorithm are shown in Fig. 4.

A. Submap Generation

Compared with the classical fixed-LiDAR, the view-angle of rotary-LiDAR varies greatly (Fig. 3). Thus, when the rotary-LiDAR visits the same location at different times, the scene expressed by the two scans may not be the same. In addition, through comprehensive analysis of Fig. 1(b) and Fig. 3, we found that due to the occlusion of the rotary-platform, a scan of the rotary-LiDAR has less integrity in scene expression

than the fixed-LiDAR. Fortunately, the submap can more fully express the appearance of the scanned scene compared to scan. Therefore, our proposed RLS-LCD algorithm adopts the place recognition strategy of submap to submap. The first step of the RLS-LCD algorithm is to generate the submap M_f of the LiDAR frame K_f according to the six-degree-of-freedom (6-DoF) pose and timestamp of each LiDAR frame. It should be emphasized that for rotary-LiDAR scan datasets, the 6-DoF pose of LiDAR frames is output by our mobile mapping algorithm suitable for rotary-LiDAR scans. In addition, the submap is similar to the local map of the classical LiDAR SLAM, but the difference is that the submap of each scan must have the same extents and point cloud density as much as possible, which helps to improve the performance of the place recognition. The detailed procedures for the submap generation are:

- 1) Adjacent LiDAR frames selection. According to the timestamp of LiDAR frames, taking K_f as the center in the frame sequence, we first select L_r (in this paper, $L_r = 5$) frames whose timestamps are less than K_f and adjacent to K_f . Similarly, L_r frames whose timestamps are greater than K_f and adjacent to K_f are then chosen. So, the submap M_f of K_f contains $2 * L_r + 1$ frames.
- 2) Submap generation. According to the 6DoF pose of the LiDAR frame, transform the point cloud of the lidar keyframe obtained in step 1) to the global frame, and we can obtain the initial submap of K_f .
- 3) Standardization of submaps. The purpose of submap standardization is to make the submap have a similar extent and point cloud density. The radius filtering on the initial submap generated in step 2) is carried out with the pose of K_f as the center, whereafter the voxel filtering is conducted on the results of radius filtering. It should be emphasized that in order to benefit the downstream tasks, we retain the maximum intensity readings within each voxel grid in voxel filtering. The submap obtained by the above processing has similar extents, and resolutions.

B. Descriptor Generation and Remapping

As shown in [30], the intensity readings returned by different surfaces are different. Thus, the LiDAR intensity readings can be employed to reveal the structure information of the surrounding environment. Inspired by the ISC algorithm [9], we construct a unique global descriptor for each submap according to the LiDAR intensity readings. This global descriptor can effectively integrate the geometric and intensity features of the submap. In addition, inspired by the fact that remapping high-dimensional data to low-dimensional in nearest neighbor search can accelerate query speed, we utilize the locality sensitive hashing (LSH) algorithm [31] to remap global descriptors to 1-dimensional to improve the efficiency of loop closure candidates retrieval in subsequent steps.

Here, we provide the principle and process for constructing global descriptors for submaps based on intensity information. The submap M_f is defined as $M_f = \{p_1, p_2, \dots, p_n\}$, where $p_k = [x_k, y_k, z_k, \eta_k]$ represents the coordinate $[x_k, y_k, z_k]$ in

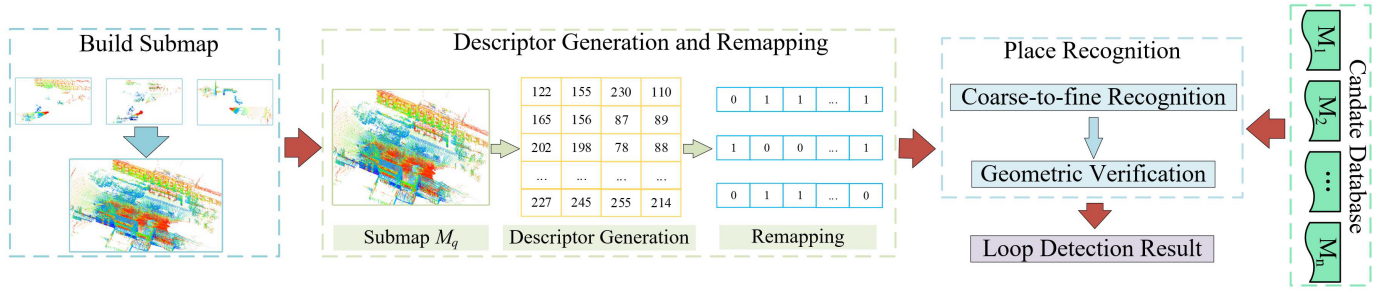


Fig. 4. The flow of the RLS-LCD algorithm.

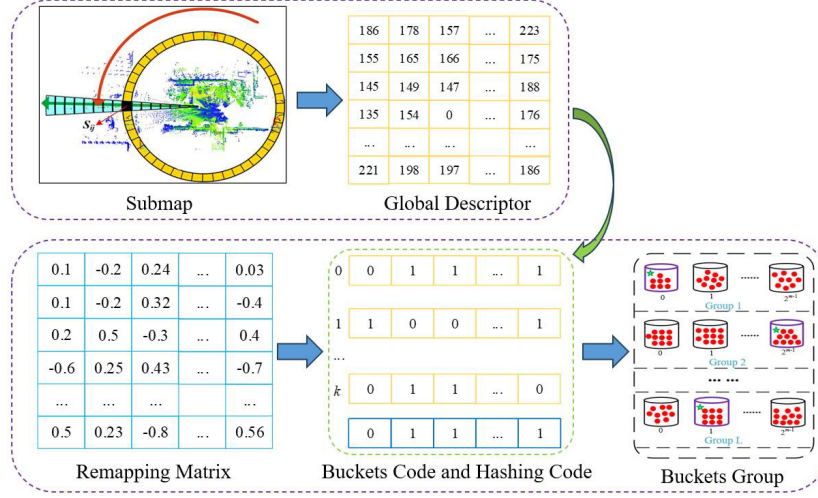


Fig. 5. The construction principle of the global descriptor and Hashing Remapping.

the global coordinate system and the intensity reading η_k of p_k . Assuming that $[x_0, y_0, z_0]$ is the LiDAR center position of the frame K_f in the global coordinate system, then each point in M_f can be transferred to the polar coordinate system in the XOY plane according to the Eq (1).

$$\begin{cases} p_k = [\rho_k, \theta_k, z_k, \eta_k] \\ \rho_k = \sqrt{(x_k - x_0)^2 + (y_k - y_0)^2} \\ \theta_k = \arctan\left(\frac{y_k - y_0}{x_k - x_0}\right) \end{cases} \quad (1)$$

Then, the middle point is divided into the subspace (as shown in Fig. 5) according to the azimuth and radius:

$$S_{ij} = \left\{ p_k \in M_f \mid \frac{i \cdot L_{max}}{N_r} \leq \rho_k < \frac{(i+1) \cdot L_{max}}{N_r}, \frac{j \cdot 2\pi}{N_s} - \pi \leq \theta_k < \frac{(j+1) \cdot 2\pi}{N_s} - \pi \right\}, \quad (2)$$

where $i \in [1, N_r]$, $j \in [1, N_s]$, $[1, N]$ the symbol denotes $\{1, 2, \dots, N\}$, and L_{max} is the maximum range of the LiDAR point. Here, K_r and K_s is the number of rows and columns in the global descriptor (in this paper, $N_r = 20$, $N_s = 120$).

Compared with the scan, the density of points in the submap is higher, so each subspace of descriptor contains a larger number of points. While recording the global information to enable the descriptor Ω to record more detailed local structure information of the submap, we take the maximum intensity value in each subspace S_{ij} as the descriptor Ω :

$$\Omega(i, j) = \max(S_{ij}). \quad (3)$$

Note that, if $S_{ij} \in \emptyset$, then $\Omega(i, j) = 0$.

After doing the above procedure, we obtain the global descriptor Ω_f of the submap M_f . To accelerate the retrieval of candidates, inspired by feature point matching in 3D reconstruction, we remap the global descriptor to lower dimensional Hamming space. Specifically, we harness the LSH algorithm to generate global remapping functions f_{bucket} and f_{hash} , then use f_{bucket} and f_{hash} to remap the global descriptors to 1-dimensional Hamming space. In principle, the essence of the remapping function is the criterion of classification, through which a variety of global descriptors can be mapped to limited categories, thereby improving the accuracy and efficiency of the recognition procedure. At the algorithm implementation level, f_{bucket} is a set $Set_{mapBucket} = \{Mat_{b_1}, \dots, Mat_{b_k}\}$ of remapping matrix generated by a (0, 1) Gaussian distribution, which consists of k matrices with m rows and N_s columns. In the same way, f_{hash} is a remapping matrix $Mat_{mapHash}$ with n rows and N_s columns generated by (0, 1) Gaussian distribution. Hashing remapping processing uses mapping functions f_{bucket} and f_{hash} to remap the global descriptors to 1-dimension as follows (Fig. 5):

- 1) Calculate the mean vector \mathbf{v} of the global descriptor Ω :

$$\mathbf{v}_{j, j \in [0, N_s]} = \frac{\sum_{i=0, j}^{N_r} \Omega(i, j)}{N_r}. \quad (4)$$

- 2) Calculate the bucket ids of the global descriptor Ω .

The matrix in $Set_{mapBucket}$ and mean vector \mathbf{v} do dot product operation according to Eq. (5) to get k binary codes of length m -bit, which is then transferred to decimal format to get k bucket ids of the descriptor Ω . Finally, k bucket ids are stored in the hash table T_{hash} . As the bucket ids between similar global descriptors are most likely to be the same, the 1-dimensional retrieval of loop closure candidates can be realized in the hash table T_{hash} through bucket ids.

- 3) Calculate the hashing code of the global descriptor Ω . The hashing code of n -bit is obtained by dot product operation of mapping matrix $Mat_{mapHash}$ and mean vector \mathbf{v} according to Eq (5). In rough filtering, hashing code can be employed to verify the similarity between candidate and query.

$$h_{r_j}(\mathbf{v}) = \begin{cases} 1, & \text{if } \mathbf{r}_j \cdot \mathbf{v} > 0 \\ 0, & \text{if } \mathbf{r}_j \cdot \mathbf{v} \leq 0 \end{cases}. \quad (5)$$

In the above expression, r_j is the j -th line of the remapping matrix.

C. Loop Closure Detection Based on Coarse-to-fine Recognition

In the RLS-LCD algorithm, we use the submap to submap place recognition strategy to reduce the influence of the view-angle change of the rotary-LiDAR. Compared to scans, submaps have richer information and stronger robust to occlusion and view-angle changes (Fig. 6). To alleviate the impact of similar scenarios on the accuracy of place recognition, we have designed a coarse-to-fine recognition strategy. To be specific, the loop closure candidates are retrieved in the hash table T_{hash} according to the bucket ids of the query first. Then, the rough filter is executed according to the similarity between the hashing code of the query and the candidate. Thereafter, the fine filter is conducted according to the similarity of global descriptors.

By leveraging the C2F retrieval technique, enhanced recall can be attained while upholding precision. For tasks related to mobile mapping, the loop closure constraints incorporated must exhibit substantial accuracy. In other words, opting for a more conservative addition of loop closure constraints is permissible, but introducing an erroneous constraint is intolerable. To address this concern, geometric validation will be employed to meticulously sift through candidate submaps. The specific process of place recognition performed by the RLS-LCD algorithm is as follows:

- 1) The initial candidates finding are based on hashing lookup. The initial candidate set $C_{ini} = \{c_1^{in}, c_2^{in}, \dots, c_n^{in}\}$ in the hash table T_{hash} is searched and generated according to the k bucket ids of the query q .
- 2) Rough filter. The Hamming distance $dist_h$ of the hashing code between the query q and the initial candidate c_i^{in} is calculated first. If $dist_h < \varepsilon$, c_i^{in} will be added to the coarse filter set $C_{coa} = \{c_1^c, c_2^c, \dots, c_n^c\}$.
- 3) Fine filter. The similarity score $\varphi(\Omega^q, \Omega^c)$ between the descriptors Ω^q and Ω_i^c of the query q and the candidate

c_i^c is computed according to Eq. (6):

$$\varphi(\Omega^q, \Omega_i^c) = \frac{1}{N_s} \sum_{j=0}^{N_s-1} \left(\frac{\mathbf{v}_j^q \cdot \mathbf{v}_j^c}{\|\mathbf{v}_j^q\| \cdot \|\mathbf{v}_j^c\|} \right). \quad (6)$$

To increase the robustness of the RLS-LCD algorithm to changes in view-angle, we calculate the similarity between all possible column-shifted descriptors Ω_l^q and the candidate descriptors Ω_i^c :

$$\Phi(\Omega^q, \Omega^c) = \max_{l=0}^n (\varphi(\Omega_l^q, \Omega^c)). \quad (7)$$

If $\Phi(\Omega^q, \Omega^c) > \tau_{th}$ (in this paper, $\tau_{th} = 0.8$), c_i^c is added to the fine filtering collection $C_{fin} = \{c_1^f, c_2^f, \dots, c_n^f\}$.

- 4) Geometric consistency verification based on submap to submap registration. The loop closure candidate obtained by intensity structure information matching uses the appearance and intensity information of the scene. In order to output loop closure constraints and ensure that the loop pairs identified using the C2F strategy have sufficient accuracy, we geometrically verify the retrieval candidate c_i^f obtained in step 3). Compared with the scan, the data amount of point cloud in the submap is larger. To improve the efficiency of geometric verification, we adopt the voxel-based registration method [32] to register the submaps of query q and candidate c_i^f . The indicator of geometric verification is the interior ratio. Specifically, we align submap M_q and submap $M_{c_i^f}$ using the registration result $T_q^{c_i}$, and then count the number of corresponding points N_{co} that satisfy $dist(p_i^q, p_j^c) < d_{inlier}$ (in this paper, $d_{inlier} = 0.2$). If $N_{co}/N_{M_q} > \tau_{inlier}$ (in this paper, $\tau_{inlier} = 0.68$), it indicates that the geometric verification has passed. Otherwise, this may be an unreliable loop pair. Here, N_{M_q} is the size of the submap M_q , LiDAR points p_i^q and p_j^c belong to the submaps M_q and $M_{c_i^f}$, respectively.

IV. EXPERIMENT EVALUATION

In this section, we validated the effectiveness of the proposed RLS-LCD algorithm in detail. In the following part, the details of the datasets and experimental settings are first introduced. Second, the influence of the parameter setting of the RLS-LCD algorithm on the performance of place recognition is discussed to determine the best parameter configuration. Third, the comparison between the performance of the proposed RLS-LCD and the state-of-the-art methods in place recognition will be demonstrated on the rotary-LiDAR dataset. Finally, the efficiency and completeness of our RLS-LCD will be evaluated. All experiments are carried out on the same platform with an Intel i7-1165G7 CPU@2.8GHz, 32GB of memory and an Ubuntu 20.04 system.

A. Dataset and Experimental Settings

The RLS-LCD algorithm is proposed mainly to improve the performance of place recognition for rotary-LiDAR scans, so we first collect a rotary-LiDAR dataset by a handled mobile mapping system (Fig. 7(a)) to verify the effectiveness of our

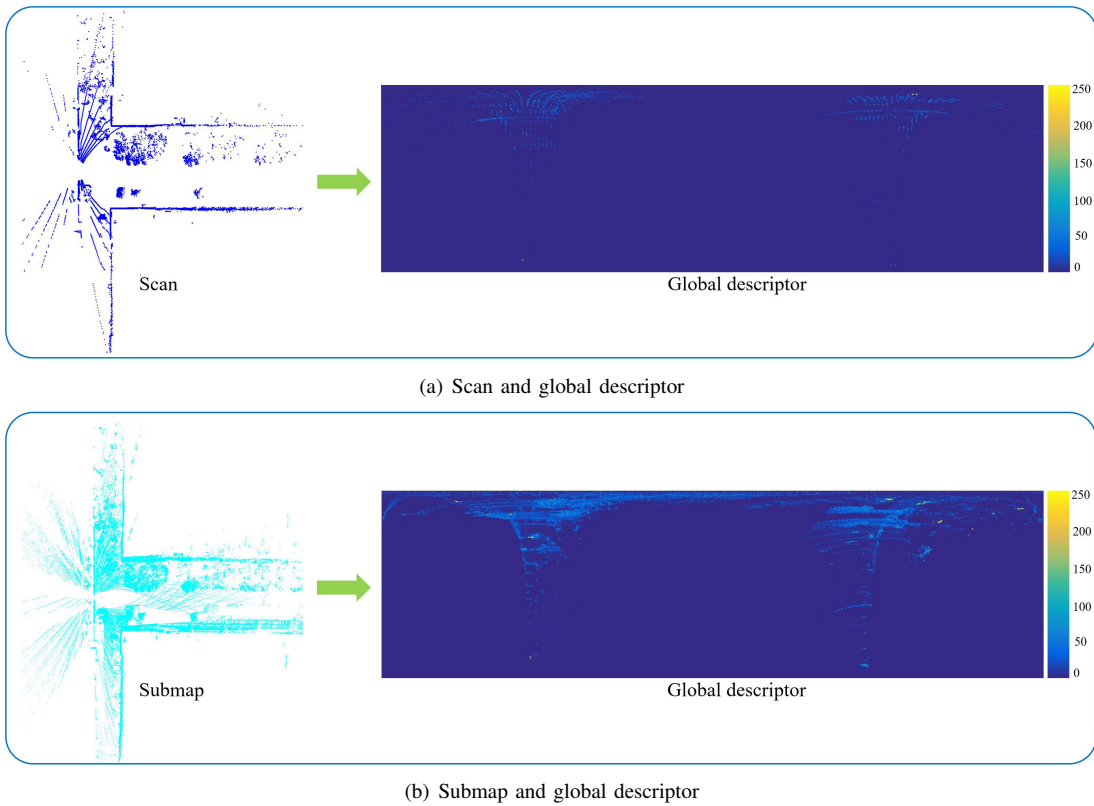


Fig. 6. Comparison of point clouds and global descriptors between scan and submap. From the graph, we can see that compared to the scan, the submap has richer information for scene representation.

method. We employed the rotary-LiDAR based on the 3D mechanical LiDAR (the 3D LiDAR information is shown in Table I), and the rotational speed of the rotary-LiDAR is 25 rpm. The rotary-LiDAR dataset includes three sequences: Sci-Tech Park, Urban Street, and Underground Parking, whose appearances are shown in Fig. 8. The details of the three LiDAR scan sequences are shown in Table II. In this testing, the developed mobile mapping algorithm is employed to output the pose of each LiDAR frame and then generate the positive pairs for the rotary-LiDAR dataset according to the pose and timestamp of each LiDAR frame. To reduce the influence of cumulative drift on generating the positive pairs, we design multiple revisits in the scanning path (Fig. 9), which can generate multiple closed loops for global optimization. Global optimization can effectively reduce the drift and ensure the reliable accuracy of the pose for each LiDAR frame. Similar to [13], the two LiDAR frames are considered a positive pair if the Euclidean distance of these two LiDAR frame poses is less than 5 meters and the timestamp interval is more than 30 seconds.

To demonstrate the advantage of the proposed RLS-LCD, three well-known algorithms based on the global descriptor, i.e. SC [4], ISC [9] and M2DP [11] and the latest method based on multi-layer contour matching (i.e. Cont2 [17]) are taken into comparison using the frequently-used metrics, recall, precision, and F_1 -score. It should be emphasized that throughout the entire experiment, the four algorithms mentioned above all used their default optimal parameter configurations, and the parameter configurations of the RLS-LCD algorithm are

TABLE I
3D LIDAR SENSOR INFORMATION

Models	Hesai XT16
Lines	16
Ranging /m	0.05-120
Ranging accuracy /cm	0.5
Vertical FoV	-15°-15°
Vertical angle resolution	2°
Horizontal angle resolution	0.18°
Data collection frequency	10HZ

TABLE II
ROTARY-LIDAR DATASET DETAILS

Dataset	Scan time /s	Trajectory length /m	Number of LiDAR frames	Number of GT loop pairs
Sci-Tech Park	541.31	727.376	2270	400
Urban Street	550.56	803.793	2468	459
Underground Parking	418.26	672.354	1788	897

shown in Table III. Considering the proposed RLS-LCD algorithm is to be integrated into the handled mobile mapping system with low computing power for real-time operation, those methods based on learning that require high computing power are not included in the comparison.

B. Parameters Setting Experiment

To obtain the best parameter configuration of the RLS-LCD algorithm, we use the control variable method to study the relationship between the setting of each parameter in



Fig. 7. Data collection.



Fig. 8. Scan scene of the rotary-LiDAR dataset.

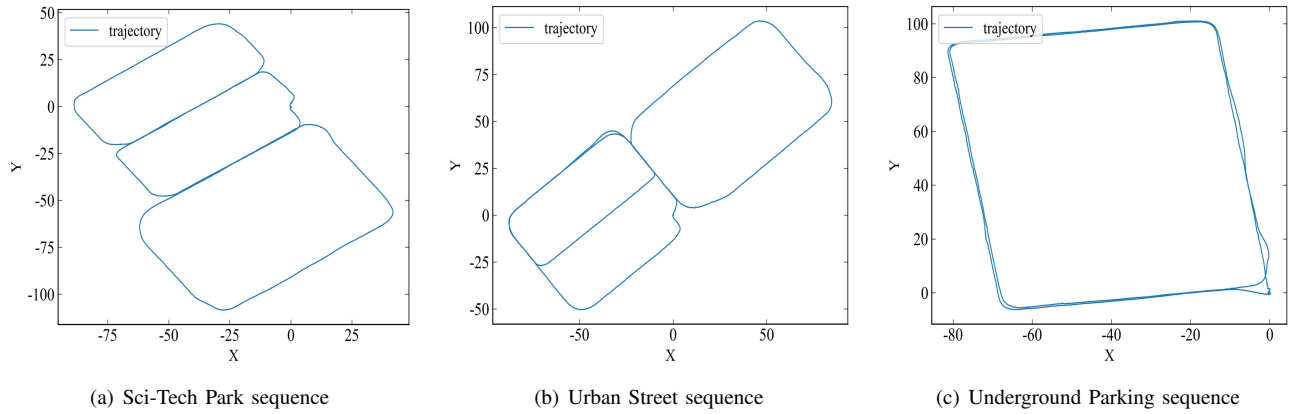


Fig. 9. Scanning trajectory for collecting rotary-LiDAR datasets.

TABLE III
PARAMETERS LIST

Parameter	Description	Default Value
N_r	Number of rows in global descriptor	20
N_s	Number of columns in global descriptor	120
k	Number of bucket group	64
m	Number of bucket bits	16
n	Number of hashing bits	64
ε	Hamming distance threshold	12

Table III and the performance of place recognition on the Sci-Tech Park sequence. The results are shown in Fig. 10. The commonly used quantitative evaluation indexes of place recognition include recall, precision, and F_1 score, in which the F_1 score takes into account both recall and precision (Eq (8)), which can reflect the performance of the algorithm more comprehensively. So we employ the F_1 score to quantitatively evaluate the performance of the RLS-LCD algorithm when we choose different parameter configurations.

$$F_1 = 2 \times \frac{P \times R}{P + R}, \quad (8)$$

where P denotes the precision and R indicates the recall.

We can see from Fig. 10(a) and Fig. 10(b) that the RLS-LCD algorithm has the best performance when $N_r = 20$ and $N_s = 120$. The reason is that, compared with scan, the point cloud density of the submap is higher and the detailed information is more abundant. Therefore, the descriptor of the high-dimensional descriptor is conducive to the expression of detailed information about the scene and can improve the performance of place recognition. For the parameters related to hashing remapping, i.e. the number of bucket group k , the number of bucket code bits m , and the number of hashing code bits n , when $k = 64$, $m = 16$, and $n = 64$, the RLS-LCD algorithm has the best performance. The setting of hamming distance threshold mainly affects the recall and precision of coarse-filtered. It can be found from Fig. 10(f) that $\varepsilon = 12$ is the best choice. The theoretical rationale behind the Hamming distance threshold is that a smaller threshold imposes more

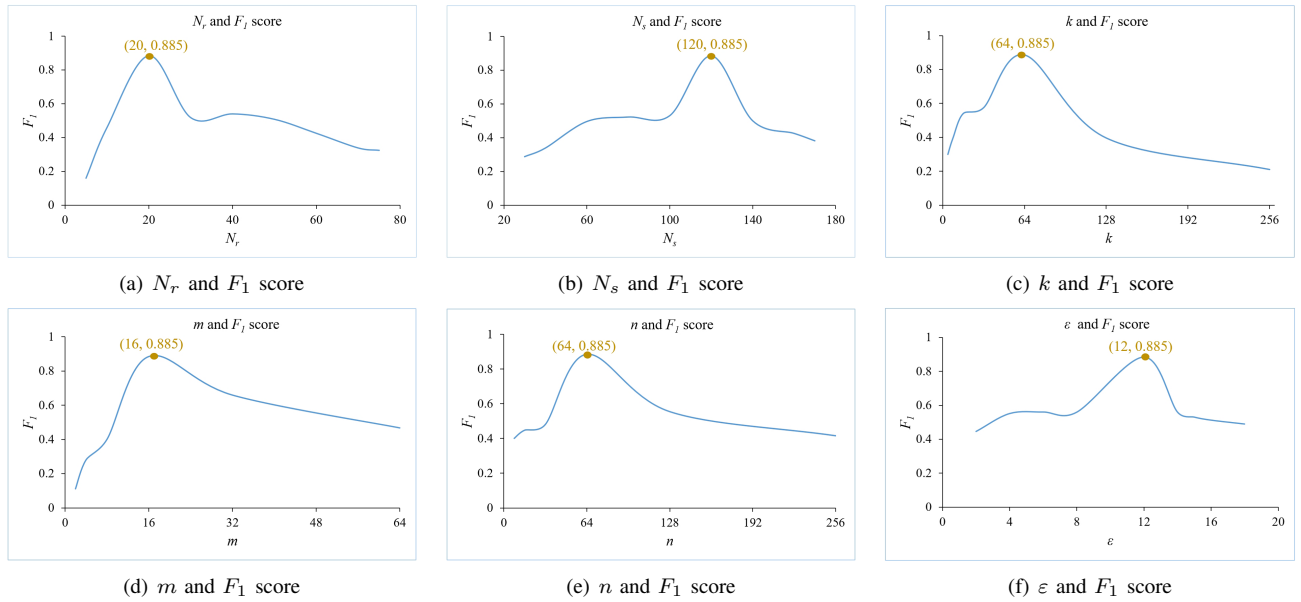


Fig. 10. The influence of Parameter setting on the performance of the RLS-LCD algorithm.

stringent conditions during coarse filtering. This configuration may yield heightened accuracy but at the expense of a notable reduction in recall. Conversely, a higher threshold enhances recall but comes at the cost of diminished precision and execution efficiency. Consequently, the selection of threshold parameters should be guided by the testing purpose and specific performance requirements, allowing for a judicious choice that optimally balances accuracy, recall, precision, and execution efficiency.

C. Effectiveness Evaluation of RLS-LCD for Rotary-LiDAR Scans

Based on the collected rotary-LiDAR dataset, the performance of the proposed RLS-LCD is compared with the SC [4], ISC [9], M2DP [11] and Contour Context (Cont2) [17]. Considering that the original SC, ISC, M2DP and Cont2 algorithms all adopt the scan to scan (S2S) place recognition strategy, we first compare the performance of the five algorithms using the S2S recognition strategy for rotary-LiDAR scans (Fig. 11 and Table IV). In addition, to verify the effectiveness of the submap-based recognition strategy in mitigating the impact of large changes in scanning view-angle of rotary-LiDAR. The performance of the five algorithms employing the M2M place recognition strategy is compared (Fig. 12 and Table V). It should be pointed out that in order to ensure fairness in testing, all five algorithms mentioned above have performed geometric consistency verification processing which is in Section III-C step 4).

From Fig. 11 and Table IV, we can see that when employing the S2S place recognition strategy, the F_1 max score of the four algorithms for rotary-LiDAR scans closed-loop recognition is less than 0.5, and the detection accuracy decreases exponentially with the increase of recall. Considering that some tasks in mobile mapping can achieve fewer revisits, such recognition performance is not satisfactory. The main reason

for this phenomenon is that the rotary-LiDAR has a large change in view-angles, and the overlap between the current scan and the historical scan near that location is very low during the revisit (Fig. 3). In addition, compared to submap, scan lacks sufficient representation of scene integrity, so the information in the global descriptor of scan is also relatively scarce (Fig. 6).

TABLE IV
 F_1 MAX SCORE BASED ON S2S RECOGNITION STRATEGY

Methods	Sci-Tech Park	Urban Street	Underground Parking	Mean
SC-S2S [4]	0.417	0.387	0.285	0.363
ISC-S2S [9]	0.479	0.417	0.416	0.436
M2DP-S2S [11]	0.401	0.351	0.282	0.345
Cont2-S2S [17]	0.477	0.476	0.333	0.429
Ours-S2S	0.466	0.479	0.435	0.449

Combining Table IV and Table V, we can find that compared with the S2S identification strategy when the M2M strategy is adopted, the performance of the all five algorithms for the place recognition of rotary-LiDAR scans has been significantly improved. Specifically, the average F_1 max scores of SC, ISC, MDP, COnt2 and RLS-LCD are improved by 126.5%, 89.7%, 113.04%, 93.9% and 97.8%, respectively. The reason here is that compared with the scan, the submap can more fully express the scene where the scan is located, that is, the submap records more detailed information about the scene, so the place recognition based on the submap will have more stability and better performance. Based on this, we can determine that the M2M strategy is effective in improving the performance of loop closure detection for rotary-LiDAR scans.

We need to focus on discussing the performance of five algorithms using the M2M strategy on the rotary-LiDAR dataset. From Table V, it can be observed that our method has the highest average score, with the Cont2 algorithm ranking second and the ISC algorithm ranking third. The similarity between our method and the ISC algorithm is that position

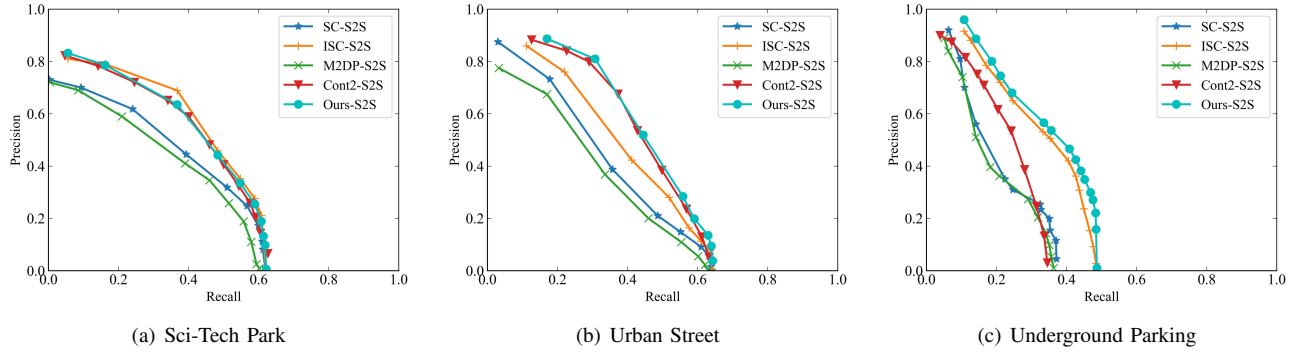


Fig. 11. Performance comparison based on S2S recognition strategy.

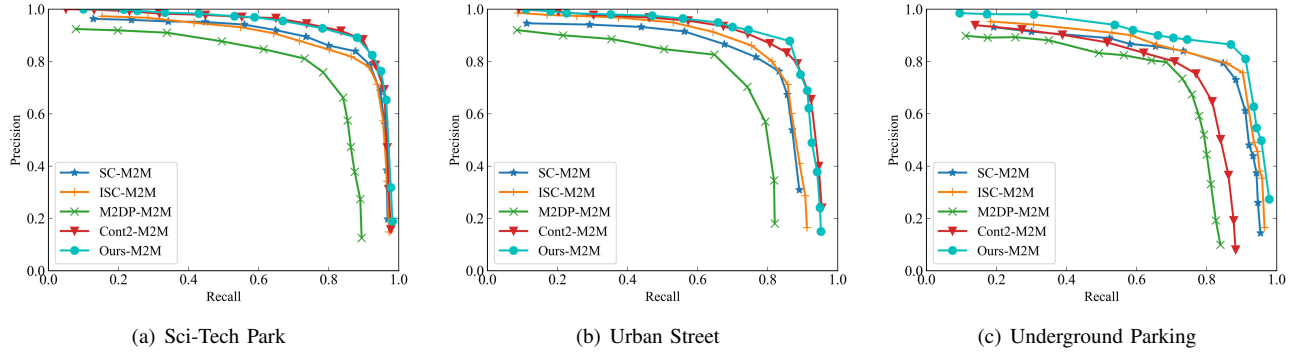


Fig. 12. Performance comparison based on M2M recognition strategy.

TABLE V
 F_1 MAX SCORE BASED ON M2M RECOGNITION STRATEGY

Methods	Sci-Tech Park	Urban Street	Underground Parking	Mean
SC-M2M [4]	0.857	0.796	0.819	0.824
ISC-M2M [9]	0.842	0.806	0.824	0.824
M2DP-M2M [11]	0.771	0.726	0.737	0.745
Cont2-S2S [17]	0.890	0.844	0.761	0.832
Ours-M2M	0.886	0.870	0.867	0.888

recognition relies on the intensity information. The difference is that we use the C2F strategy for loop closure candidates retrieval, which is reliable in similar structural scenarios. The ISC algorithm uses a search method similar to exhaustive traversal. Exhaustive traversal is inefficient and susceptible to the influence of similar structures. Therefore, aggregating the test results in all three scenarios, our method outperforms the ISC algorithm in terms of performance.

Subsequently, we delve into the performance distinctions between the Ours and Cont2 methodologies. The specific processing of the Cont2 algorithm is to first convert each LiDAR frame to the BEV frame, then extract the contour of the point cloud at multiple heights and model each contour using a 2D Gaussian distribution. The place recognition processing adopts contour matching measured in L2 distance. The characteristic of rotary-LiDAR scanning is that the view-angle changes greatly, particularly in roll and pitch. Therefore, preceding the application of the Cont2 algorithm, it becomes imperative to translate the submaps from each LiDAR frame to

the global frame and undergo decentralization. However, when the scanning distance is too long, the drift in the Z direction may affect the uniqueness of contour extraction. Furthermore, the Cont2 algorithm, which relies on contour matching, may encounter challenges when dealing with structurally similar scenes. The performance analysis of Cont2 and our method in three sequences is as follows:

- 1) Sci-Tech Park sequence. The structural types of the Sci-Tech Park are diverse, and the drift of the scanning trajectory is small, so Cont2 has reliable performance (Fig. 12(a)). Moreover, given the ample intensity information available in this scenario, the performance of our method closely aligns with that of the Cont2 method. A meticulous examination of subgraphs A and B in Fig. 13(a) and Fig. 13(b) reveals a comparable performance between our method and the Cont2 method.
- 2) Urban Street sequence. The scene structure in this sequence is relatively single in type compared to the Sci-Tech Park scene, but it has rich intensity information. We can see from Fig. 12(b) that the performance of the Cont2 method is weaker than our method.
- 3) Underground Parking sequence. There are many similar building structures, but the cars, iron gates, pipelines and lane lines in the scene have rich strength information. Upon careful analysis of subgraphs A and B in Fig. 13(c) and Fig. 13(d), we can see that Cont2 cannot detect loop pairs in many regions. However, our method employs intensity information and the C2F retrieval strategy, resulting in more stable performance in the Underground Parking scenario.

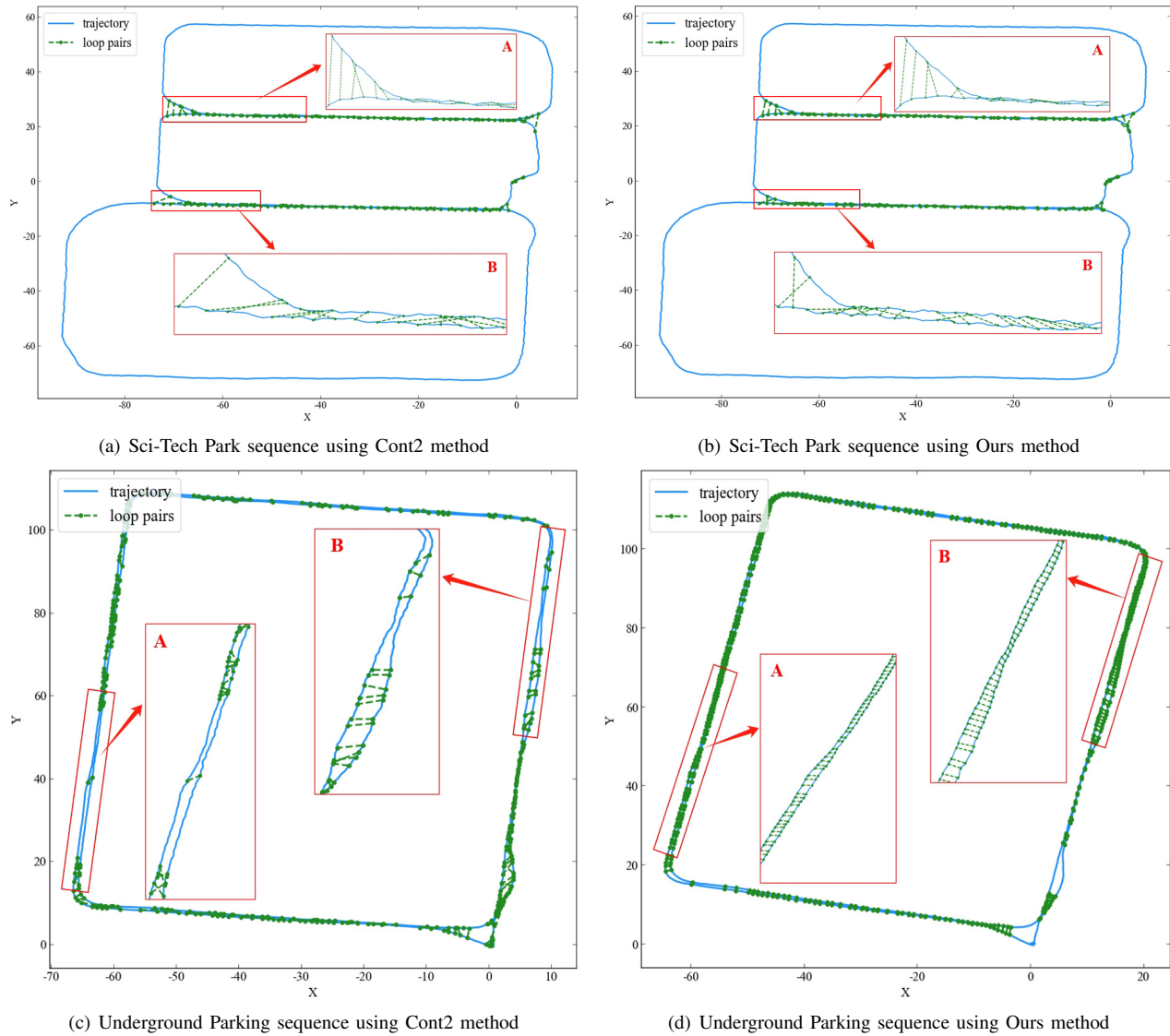


Fig. 13. The place recognition results using Cont2 and Ours methods in two representative scenarios. We can see that our method performs similarly to the Cont2 method in the Sci-Tech Park scenario. In the Underground Parking scenario, the performance of the Cont2 method is inferior to our method due to the influence of a single structural type.

D. Effectiveness Evaluation of Coarse-to-fine Place Recognition Strategy

We have modified the closed-loop candidate retrieval method of the RLS-LCD algorithm to the traditional kd-tree method [33] method and our designed coarse-to-fine (C2F) method, respectively, for place recognition testing on rotary-LiDAR datasets to verify the effectiveness of the C2F strategy. It should be pointed out that in the above comparative tests, we focus on the effectiveness of the C2F strategies, so all methods involved in the comparison will not perform geometric consistency validation. The outcomes of the conducted tests are presented in Fig. 14. Through a comparative analysis of the results obtained from the Ours-KDTree and Ours-C2F methodologies, we discerned that the adoption of the C2F recognition strategy leads to a noticeable enhancement in place recognition performance. The distinct contribution of the C2F strategy primarily manifests in precision gains. This can be attributed to the fact that Ours-KDTree employs a singular

filtering mechanism, whereas the C2F strategy employs a multi-layer filtering mechanism in its candidate search process.

The main advantage of the multi-layer filtering mechanism is that it can suppress the impact of similar structures of scene on place recognition accuracy. Based on the analysis of the characteristics employed for testing scenarios, there are a large number of similar appearance structures in both Urban Street and Underground Parking. In such scenarios, the advantages of Ours-C2F method will be fully reflected. This viewpoint is confirmed in Fig. 14(b) and Fig. 14(c). At the principle level, the underlying reason is that compared to the KD-Tree method, which only performs candidate filtering based on the similarity of the mean vector of the global descriptor, our C2F strategy requires three layers of filtering, namely: candidate set search based on bucket id, coarse filtering in Hamming space, and fine filtering based on global descriptor similarity.

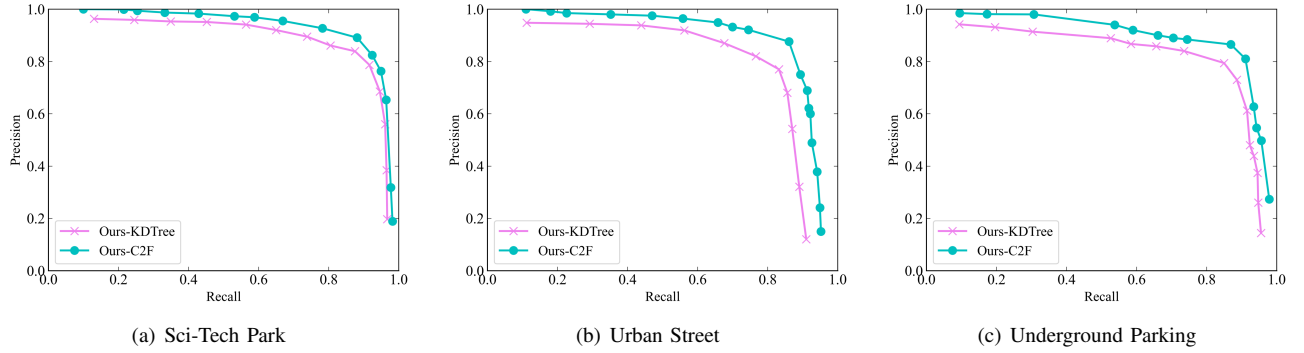


Fig. 14. Effectiveness analysis of coarse-to-fine recognition strategy. We used the traditional KD-Tree method and our designed coarse-to-fine strategy for rotary-LiDAR dataset, and found that using the coarse-to-fine recognition strategy is effective in improving the performance of closed-loop detection.

TABLE VI
EFFICIENCY COMPARISON

Methods	Descriptor generation /ms	Loop closure candidates retrieval /ms	Sum /ms
SC [4]	9.28	9.43	18.71
ISC [9]	9.12	8.42	17.54
M2DP [11]	8.89	8.28	17.17
Cont2 [17]	11.16	14.50	25.66
Ours	12.4	13.5	25.9

E. Efficiency and Completeness Evaluation

For place recognition, the computational complexity is an essential indicator to assess the merit of an algorithm, which is especially true for the handled mobile mapping system. In this part, we evaluate and report the time cost of the SC, ISC, M2DP, Cont2 and RLS-LCD algorithm in the Sci-Tech Park sequence (Table VI). As can be seen from Table VI, the total average time spent on descriptors generation and loop closure candidates retrieval for SC, ISC, MDP, Cont2 and RLS-LCD algorithms is 18.71 ms, 17.54 ms, 17.17 ms, 25.66ms and 25.9 ms, respectively. Therefore, the RLS-LCD algorithm has the highest time cost because its process is more complex. However, according to Fig. 12, our comprehensive evaluation found that a slight loss in efficiency is worthwhile when the recall rate and accuracy are significantly improved. In addition, our developed mobile mapping system has a time interval of 100ms between LiDAR frames (10HZ). Although RLS-LCD is slower than the other three algorithms, it can still meet the requirements of real-time processing.

To evaluate the completeness, the RLS-LCD algorithm is integrated into our handheld mobile mapping system (Fig. 1(b)). The total station is used to measure four control points (Fig. 7(b)). In the Street scene, three points are employed for coordinate transformation and the remaining one is set as a checkpoint (Fig. 15(a)). In the Parking scene, four points are employed for coordinate transformation and two points are used as checkpoints (Fig. 15(b)), of which 4 are employed for coordinate transformation and two control points are exploited as checkpoints. Table VII shows the accuracy of the checkpoints for the two mobile mapping with and without the RLS-LCD algorithm. Here, we measure the absolute accuracy of mobile mapping by calculating the difference (ΔX , ΔY and ΔZ), between the coordinates of the checkpoint output from

the total station measurement and the mobile mapping system. It can be seen that the accuracy of mobile mapping using RLS-LCD can be significantly improved. Therefore, based on comprehensive efficiency and effectiveness evaluation, our proposed RLS-LCD algorithm is capable of performing place recognition tasks in a mobile mapping system integrating rotary-LiDAR.

TABLE VII

ACCURACY OF CHECKPOINTS IN COMPLETENESS EVALUATION

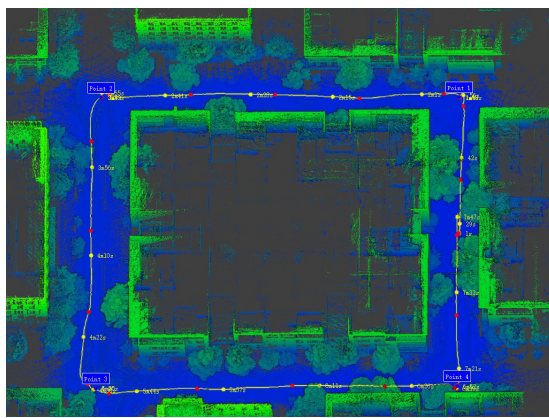
Sequence	checkpoint ID	ΔX	ΔY	ΔZ
Street	Point 3	1.22	0.95	1.12
	Point 3	1.61	1.12	1.24
Parking	Point 6	2.33	0.58	1.36
	Point 3	2.42	1.35	2.18
Street (No RLS-LCD)	Point 3	2.66	2.45	3.12
Parking (No RLS-LCD)	Point 6	3.56	2.89	3.23

V. CONCLUSION

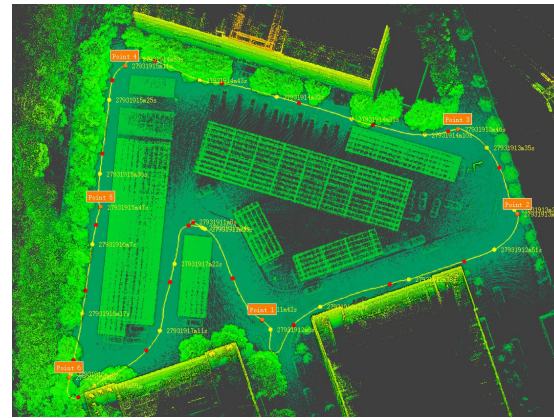
In this paper, we proposed an effective LCD method for rotary-LiDAR scans. Compared with the LCD methods for fixed-LiDAR scans, the proposed RLS-LCD algorithm could effectively address the large changes in scanning view-angle by designing a submap-based place recognition method and a coarse-to-fine place recognition strategy. A series of experiments on three scan sequences demonstrated that our proposed method achieved state-of-the-art place recognition performance for rotary-LiDAR scans. Moreover, the efficiency and integrity evaluation showed that our method could meet the requirement of real-time place recognition of the mobile mapping system that integrated rotary-LiDAR. However, our method relies on the intensity reading of LiDAR and does not consider the impact of dynamic objects on place recognition. Therefore, our future research direction is to improve the robustness of place recognition in areas with insufficient intensity information and reduce the adverse impact of dynamic objects.

REFERENCES

- [1] H. Yin, L. Tang, X. Ding, Y. Wang, and R. Xiong, "Locnet: Global localization in 3D point clouds for mobile vehicles," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 728–733, IEEE, 2018.



(a) Street (Point 3 is checkpoint)



(b) Parking lot (Point 3 and Point 6 is checkpoint)

Fig. 15. Mapping results of completeness evaluation.

- [2] S. Li, G. Li, L. Wang, and Y. Qin, "SLAM integrated mobile mapping system in complex urban environments," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 166, pp. 316–332, 2020.
- [3] A. Angeli, D. Filliat, S. Doncieux, and J.-A. Meyer, "Fast and incremental method for loop-closure detection using bags of visual words," *IEEE transactions on robotics*, vol. 24, no. 5, pp. 1027–1037, 2008.
- [4] G. Kim and A. Kim, "Scan context: Egocentric spatial descriptor for place recognition within 3D point cloud map," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4802–4809, IEEE, 2018.
- [5] Y. Latif, C. Cadena, and J. Neira, "Robust loop closing over time for pose graph SLAM," *The International Journal of Robotics Research*, vol. 32, no. 14, pp. 1611–1626, 2013.
- [6] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *2009 IEEE international conference on robotics and automation*, pp. 3212–3217, IEEE, 2009.
- [7] J. Ma, X. Ye, H. Zhou, X. Mei, and F. Fan, "Loop-closure detection using local relative orientation matching," *IEEE Transactions on Intelligent Transportation Systems*, 2021.
- [8] J. Guo, P. V. Borges, C. Park, and A. Gawel, "Local descriptor for robust place recognition using lidar intensity," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1470–1477, 2019.
- [9] H. Wang, C. Wang, and L. Xie, "Intensity scan context: Coding intensity and geometry relations for loop closure detection," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2095–2101, IEEE, 2020.
- [10] M. Weinmann, B. Jutzi, and C. Mallet, "Semantic 3D scene interpretation: A framework combining optimal neighborhood size selection with relevant features," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 2, no. 3, p. 181, 2014.
- [11] L. He, X. Wang, and H. Zhang, "M2dp: A novel 3D point cloud descriptor and its application in loop closure detection," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 231–237, IEEE, 2016.
- [12] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "Netvlad: Cnn architecture for weakly supervised place recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5297–5307, 2016.
- [13] X. Kong, X. Yang, G. Zhai, X. Zhao, X. Zeng, M. Wang, Y. Liu, W. Li, and F. Wen, "Semantic graph based place recognition for 3D point clouds," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 8216–8223, IEEE, 2020.
- [14] L. Li, X. Kong, X. Zhao, T. Huang, W. Li, F. Wen, H. Zhang, and Y. Liu, "Ssc: Semantic scan context for large-scale place recognition," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2092–2099, IEEE, 2021.
- [15] H. Xiang, W. Shi, W. Fan, P. Chen, S. Bao, and M. Nie, "Fastled: A fast and compact loop closure detection approach using 3D point cloud for indoor mobile mapping," *International Journal of Applied Earth Observation and Geoinformation*, vol. 102, p. 102430, 2021.
- [16] D. Cattaneo, M. Vaghi, and A. Valada, "Lcdnet: Deep loop closure detection and point cloud registration for LiDAR SLAM," *IEEE Transactions on Robotics*, 2022.
- [17] B. Jiang and S. Shen, "Contour context: Abstract structural distribution for 3D LiDAR loop detection and metric pose estimation," *arXiv preprint arXiv:2302.06149*, 2023.
- [18] Z. GeoSLAM, "Go—GeoSLAM."
- [19] Z. W. RigelSLAM, "RigelSLAM."
- [20] X. Lixel L1, "Lixel L1."
- [21] S. Karam, V. Lehtola, and G. Vosselman, "Simple loop closing for continuous 6DOF LIDAR&IMU graph SLAM with planar features for indoor environments," *ISPRS journal of photogrammetry and remote sensing*, vol. 181, pp. 413–426, 2021.
- [22] M. Bosse and R. Zlot, "Place recognition using keypoint voting in large 3D lidar datasets," in *2013 IEEE International Conference on Robotics and Automation*, pp. 2677–2684, IEEE, 2013.
- [23] T. Shan, B. Englot, F. Duarte, C. Ratti, and D. Rus, "Robust place recognition using an imaging lidar," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5469–5475, IEEE, 2021.
- [24] R. Dubé, D. Dugas, E. Stumm, J. Nieto, R. Siegwart, and C. Cadena, "Segmatch: Segment based place recognition in 3D point clouds," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5266–5272, IEEE, 2017.
- [25] M. A. Uy and G. H. Lee, "Pointnetvlad: Deep point cloud based retrieval for large-scale place recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4470–4479, 2018.
- [26] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3D classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 652–660, 2017.
- [27] Z. Liu, S. Zhou, C. Suo, P. Yin, W. Chen, H. Wang, H. Li, and Y.-H. Liu, "LPD-Net: 3D point cloud learning for large-scale place recognition and environment analysis," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2831–2840, 2019.
- [28] K. Vidanapathirana, P. Moghadam, B. Harwood, M. Zhao, S. Sridharan, and C. Fookes, "Locus: Lidar-based place recognition using spatiotemporal higher-order pooling," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5075–5081, IEEE, 2021.
- [29] J. Cheng, C. Leng, J. Wu, H. Cui, and H. Lu, "Fast and accurate image matching with cascade hashing for 3D reconstruction," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–8, 2014.
- [30] A. G. Kashani, M. J. Olsen, C. E. Parrish, and N. Wilson, "A review of lidar radiometric processing: From ad hoc intensity correction to rigorous radiometric calibration," *Sensors*, vol. 15, no. 11, pp. 28099–28128, 2015.
- [31] M. S. Charikar, "Similarity estimation techniques from rounding algorithms," in *Proceedings of the thirty-fourth annual ACM symposium on Theory of computing*, pp. 380–388, 2002.
- [32] K. Koide, M. Yokozuka, S. Oishi, and A. Banno, "Voxelized gicp for fast and accurate 3D point cloud registration," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 11054–11059, IEEE, 2021.
- [33] J. L. Blanco and P. K. Rai, "nanoflann: a c++ header-only fork of FLANN, a library for nearest neighbor (nn) with kd-trees," 2014.



Qiyuan Zhang received a master's degree from the Xi'an University of Science and Technology, Xi'an, China, in 2020. He is currently pursuing a Ph.D. degree with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China. His research interests include image matching, simultaneous localization and mapping.



Xiaonan Wang received a doctorate from Wuhan University in 2014. He currently works in Wuhan ZG Automation Technology Co., Ltd. His research interests include high-precision close-range 3D reconstruction and industrial automatic detection.



Shunyi Zheng received the Post-Doctorate from the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China, in 2002. He is currently a Professor with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China. His research interests include remote sensing data processing, image matching, and three-dimensional reconstruction.



Rui Li received a Bachelor's degree from the School of Automation Science and Engineering, South China University of Technology, Guangzhou, China, in 2019 and a Master's degree from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2021. He is pursuing a Ph.D. degree with the Intelligent Control & Smart Energy (ICSE) Research Group, School of Engineering, University of Warwick, Coventry, UK. He has authored more than 20 peer-reviewed

articles in international scientific journals such as ISPRS P&RS, IEEE TGRS, PR, APEN, ECM and Energy. His research interests lie in trans-disciplinary applications of deep learning, especially for remote sensing, computer vision and renewable energy. He was honoured with the U.V. Helava Award (Best Paper 2022) from the International Society for Photogrammetry and Remote Sensing for his work on Vision-Transformer-based semantic segmentation published in ISPRS P&RS.



Xiqi Wang received a master's degree from the Xi'an University of Science and Technology, Xi'an, China, in 2017. He is currently pursuing a Ph.D. degree with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China. His research interests include deep learning and simultaneous localization and mapping.



Yuan He received a doctorate from Wuhan University in 2020. He currently works in Wuhan ZG Automation Technology Co., Ltd. His research interests include 3D reconstruction and mobile mapping.