2023 IEEE International Conference on Robotics and Automation (ICRA)
Workshop on Active Methods in Autonomous Navigation
29 May – 2 June, 2023, London, UK

Visual-based and Lidar-based SLAM Study for Outdoor Environment

Siti Sofiah Binti Mohd Radzi [1], Siti Sarah Binti Md. Sallah [1], Nazlee Azmeer bin Massuan [1], Lee Ming Yi [1], Qamarul Aiman Bin Tajul Ariffin [1], Bayhaqi Bin Mohd Jailani[1], and Hon Hock Woon [1]

Abstract— **Malaysia's palm oil plantation industry is labor intensive. It is estimated that there were 505,972 workers in 2012 and this number consists mainly of foreigners of about 76.5% and local 23.5%. To address the labor-intensive nature of the industry and reduce costs, mechanization and automation have emerged as potential solutions. However, the dense foliage of matured oil palm trees, along with their overlap with neighboring trees, obstructs crucial information beneath the canopy. Aerial mapping or Google map images cannot reliably provide the necessary information for unmanned ground vehicles (UGVs) to navigate the under-canopy environment. Furthermore, GPS signals may be degraded or inaccessible under the canopy. Consequently, researchers have developed Vision-based and Lidar-based navigation methods, specifically tailored for GPS-denied environments, such as Simultaneous Localization and Mapping (SLAM). This survey paper aims to explore the current research on Visual-based and Lidar-based navigation in outdoor environment without relying on GPS sensor. Experiments are carried out in outdoor environment to simulate situations without GPS coverage. The resulting maps generated from these methods were then qualitatively analyzed.**

I. INTRODUCTION

The robotics community has witnessed a significant rise in outdoor applications across various sectors such as agriculture, search and rescue, mining, defense, environment monitoring, and planetary exploration. In the agricultural sector, where productivity relies on efficiency, reliability, and precision, reducing the need for manual labor is crucial as it constitutes a major portion of field operation costs. For example, Malaysia's labor-intensive oil palm plantation industry employed approximately 505,972 workers in 2012, with foreigners accounting for 76.5% and locals 23.5%.[1] Mechanization and automation are crucial for reducing labor costs and maintaining competitiveness in agriculture. However, the agricultural environment is challenging, with dynamic and unstructured conditions. Accurate mapping and localization are necessary for autonomous navigation in such environments. Traditional GPS signals may be limited or unavailable, requiring the use of Simultaneous Localization and Mapping (SLAM) techniques. SLAM compensates for GPS limitations by utilizing sensors like stereoscopic cameras or Lidar, which are combined with secondary sensors like IMUs, wheel odometry, or GPS to maintain the 3D spatial structure of SLAM. The combination of multiple sensor inputs enables vehicle path planning, obstacle avoidance, and manipulation. [2]

II. SLAM ALGORITHMS

SLAM (Simultaneous Localization and Mapping) plays a significant role in autonomous vehicles and robot decision making. It is divided into two main types [3]: visual SLAM and Lidar-Based SLAM. Visual SLAM, or VSLAM, uses cameras and image sensors to capture data, offering advantages such as rich information, affordability, lightweight design, and small size. However, it is sensitive to lighting conditions. On the other hand, Lidar-Based SLAM primarily relies on laser or distance sensors, providing greater precision and being unaffected by lighting conditions. Nevertheless, high-resolution lidar sensors can be expensive. Both types of SLAM contribute to navigation and decision-making processes by enabling simultaneous localization and mapping in various applications.

The main objective of a 3D SLAM system is to determine the robot's position and orientation while creating a map of the environment using sensor data. To build the map, classic SLAM techniques use front-end odometry, which relates subsequent sensor scans, such as lidar or stereo camera point clouds. The typical method for relating these scans is iterative closest point (ICP), but it can be computationally expensive when dealing with a large number of points. Another approach is feature-matching, which reduces computation costs. However, both methods still result in some accumulated error over time, leading to inaccurate odometry. To address this, back-end optimization is used. There are two types of back-end optimization methods: filtering approaches like Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), or particle filter, which correct the robot's state in real-time as new measurements are available, and graph optimization approaches, which optimize the full robot trajectories using a complete set of measurements to enhance odometry accuracy. The methods used in graph optimization are for example such as iSAM (Incremental Smoothing And Mapping) [4], GTSAM (Georgia Tech Smoothing and Mapping) [5], and G2O(General Graph Optimization) [6]. In SLAM algorithm, the loop closure detection is added to detect whether the path has formed a loopback. This is because the current trajectory accumulates error by time and form a significant trajectory drift. Loop closure detection eliminates drifts by eliminating visited locations to create a more accurate map. During each mapping session, global loop closure detection is used to figure out when the robot goes back to a previous map[7]. Global loop closure is important because it can fix errors that build up as the robot moves through the environment. Local loop closure, on the other hand, can make the map more consistent and complete in a given area. A user may not revisit to a particular point in the scene, which makes global loop closure impossible. Therefore, local loop closures are used to refine the initial pose estimate [8]. There are several SLAM methods, for example, RTABMAP [9], ORB-SLAM3 [10], LEGO-LOAM [11], LIO-SAM [12, 13], LVI-SAM [14] and

*Research is supported by Malaysian Ministry of Science, Technology and Innovation (MOSTI) under the strategic research grant (SRF).[1] Authors are with MIMOS Berhad Malaysia under Advanced Intelligence Lab. (sofiah.radzi@mimos.my, sarah.sallah@mimos.my, nazlee.massuan@mimos.my,mingyi.lee@mimos.my,aiman.ariffin@mimos.my, bayhaqi.jailani@mimos.my, hockwoon.hon@mimos.my)

others. In this study, we have compared two SLAM (Simultaneous Localization and Mapping) algorithms: RTABMAP and LIO-SAM using a combination of 3D Lidar, stereo camera, and IMU (Inertial Measurement Unit) as inputs. RTABMAP has the capability to use either Lidar or visual camera input for its odometry node, whereas LIO-SAM specifically utilizes Lidar and IMU data. For this research, we aimed to investigate and compare the qualitative differences in map accuracy between these two algorithms, despite both of them incorporating Lidar as one of the input sensors.

A. *RTABMAP*

RTAB-MAP (Real-Time Appearance-Based Mapping) [9, 15] is a graph-based SLAM approach widely used in mobile robot navigation within the Robot Operating System (ROS). It supports various data types such as RGB-D, stereo, and LiDAR, and is versatile in handling mixed modalities. RTAB-MAP can work with different odometry approaches, including visual, LiDAR, or wheel-based methods. For visual odometry, it uses Frame-To-Map (F2M) or Frame-To-Frame (F2F) techniques, while LiDAR odometry involves Scan-to-Map (S2M) or Scan-to-Scan (S2S) methods. The latter processes 3D point clouds from LiDAR scans, down-samples them, calculates normals, and applies the Iterative Closest Point (ICP) algorithm to estimate the transformation between fixed and moving point clouds. RTAB-MAP incorporates an appearance-based loop closure detector using a bag-of-words approach, which helps identify whether a new image corresponds to a previous or new location[15].When a loop closure hypothesis is accepted, a new constraint is added to the graph of the map, whereupon a graph optimizer minimizes the errors in the map. When a loop closure occurs, the global map should be re-assembled according to all new optimized poses for all nodes in the map's graph. For loop closure detection, the current bag-of-words approach is dependent on a camera, meaning that a camera is always required even if LIDAR is used for odometry.

B. *LIO SAM*

LIDAR is commonly utilized alongside other sensors, such as IMU and GPS, to estimate the state and create map. There are two primary design approaches for Lidar mapping that involve sensor fusion: loosely-coupled fusion and tightly-coupled fusion. In the tightly-coupled approach, the pre-integrated IMU measurements are typically leveraged to rectify the skewed point clouds. Lidar Inertial Odometry, a tightly-coupled method, employs all IMU data directly and optionally incorporates a global positioning system (GPS) to correct height errors. To achieve an accurate map over large outdoor areas, sensor fusion between LIDAR/IMU and a high-precision GPS sensor is essential.[16]. LIO-SAM formulates lidar inertial odometry on a factor graph suitable for multi-sensor fusion and global optimization [12]. In LIO-SAM four types of factors are:(a) IMU pre-integration factors, (b) lidar odometry factors, (c) GPS factors, and (d) loop closure factors. It estimates states utilizing tightly coupled IMU integration and Lidar odometry via factor graph optimization. In Lidar odometry factor, a feature extraction is performed on new lidar scan. Edge and planar features are extracted by evaluating the roughness of points over a local region. Instead of using every lidar for computing and adding factors, the LIO SAM adopt the keyframe selection. The factor graph is optimized upon the insertion of a new node using incremental smoothing and mapping with the Bayes tree (iSAM2) [17].Warku et. al [13] use a Velodyne (VLP-16) and Xsens MTI-G-700 to obtain data from inertial measurement devices to create a 3D mapping of the indoor and outdoor environment and visualize the created map using the LIO-SAM method. The map optimization optimizes lidar odometry factor and GPS factor. This factor graph is maintained consistently throughout the whole test. The IMU pre-integration optimizes IMU and the lidar odometry factor and estimates the IMU bias. This factor graph is periodically reset and guarantees real-time odometry estimation with frequency IMU. Fig. 1 illustrates the factor-graph structure in LIO-SAM flow diagram. Loop closure detected by Euclidean distance radius search and ICP registration significantly degrades real-time performance [18].

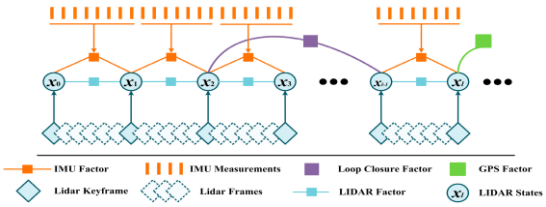
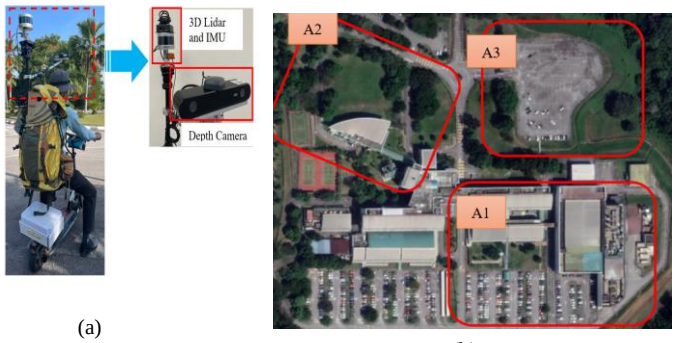


Figure 1 Factor-graph structure of LIO-SAM[12]

III. EXPERIMENTAL SETUP AND RESULTS

A. *Experimental Setup*

Our experimental site is located at the MIMOS campus in Kuala Lumpur, Malaysia. The total size of the MIMOS site is approximately 133288 m^2. The custom hardware setup with Scooter-based is shown in Figure 2(a). A ROS melodic operating system and equipped with sensors such as 3D Lidar VLP-16, Xsens IMU 710 with external receiver and a unit depth camera ZED 2i. The speed of the scooter while recording is 12km/h. The analysis of the 3D point cloud produced from MIMOS Berhad area is divided into three subareas; A1, A2 and A3 as shown in Fig. 2(b). A1 has more features, mainly parking lots. A2 contains more challenging roads (elevations, slopes, and bumps), and A3 is mostly consists of unoccupied parking lots. The user drives the scooter around MIMOS Berhad in



(a) (b)

Figure 2. (a) Experimental setup for hand-held devices based on motorized scooter. (b) Top view of the image captured by Google Earth. The datasets are collected in 3 different subareas of MIMOS: A1, A2, A3

order to collect the dataset, then pivots back to the starting locations.

B. Results

For qualitative evaluation, RTABMAP and LIO-SAM were utilized to generate a comprehensive 3D map of the environment. To evaluate loop closure detection, markers were placed in specific areas and compared using both methods. In cases where a false loop closure was identified, it was classified as a false positive, indicating incorrect loop closure detection. Conversely, if no loop closure was detected in a similar previously visited location, it was considered a false negative, indicating a missed loop

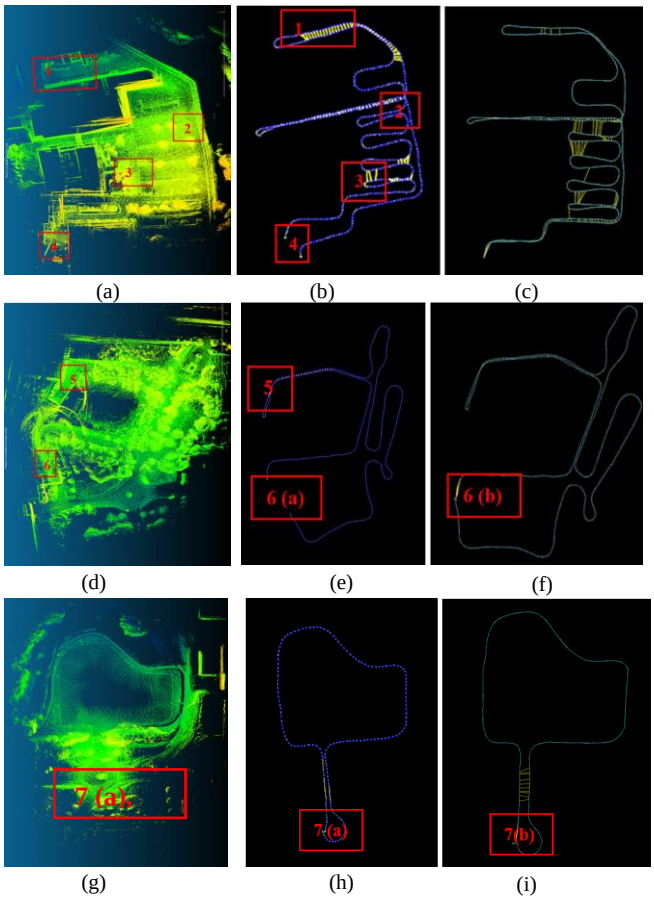


Figure 3 (a) Top view of A1 RTABMAP 3D mapping (b)RTABMAP Mapping trajectory of A1 (c) LIO SAM Mapping trajectory of AI (d) Top view of A2 RTABMAP 3D mapping (e)RTABMAP Mapping trajectory of A2 (f) LIO SAM Mapping trajectory of A2 (g) Top view of A3 RTABMAP 3D mapping (h)RTABMAP Mapping trajectory of A3 (i) LIO SAM Mapping trajectory of A3

closure.Loop closure detection can be performed through two methods: local loop closure and global loop closure. Local loop closures are focused on detecting and correcting loop closures within a restricted spatial and temporal range, typically centered around the robot's current pose or trajectory segment. On the other hand, global loop closures encompass the detection and correction of loop closures over the entire trajectory or a significant portion of it, aiming to achieve alignment and optimization at a global scale. By employing these methods, RTABMAP and LIO-SAM aim to identify and rectify loop closures in order to enhance map accuracy and maintain consistent localization throughout the mapping process.

Fig. 3(a) shows the top view of generate map of A1 using

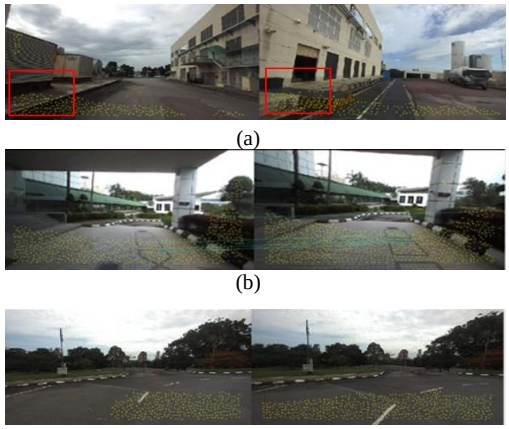


Figure 4 . (a) RTABMAP for A1: Local loop closure detected due to identical objects in red box. The yellow dotted are inliers found more than one time in the image. It is same objects' features (drain) found but at different location (b) RTABMAP for A2: No global closed loop is detected, even if the start and end points are at the same location. This is using the default parameters (c) RTABMAP for A3:No loop close detected

RTABMAP, Fig. 3(b) shows the trajectory of A1 with RTABMAP, while Fig. 3(c) shows the trajectory of A1 with LIO-SAM. Based on the observations made at marker 1, the result indicates the presence of false positive local loop closure detection in the correct area. However, in Fig. 4(a), the RTABMAP database viewer shows loop closure detection between identical objects (specifically, a drain) but actually at a different location. At marker 2, there is a false negative detection where the map fails to close a loop due to insufficient inliers in the camera image or issues with image illumination. Similarly, at marker 3, there is an incorrect local loop closure as the color of the identical object differs. Moreover, at marker 4, the start and end images in map A1 do not achieve global loop closure because of a lack of information about the U-turn, resulting in map duplication.

Fig. 3(d) provides a top view map of A2 using RTABMAP while Fig. 3(e) shows A2 with trajectory of RTABMAP and Fig. 3(f) illustrate the mapping trajectory of LIO SAM in A2. Marker 5 represents a local loop closure, indicating a successful detection. On the other hand, marker 6(a) indicates the absence of loop closure, which is considered a false negative detection. The first and last camera photos displayed in the RTABMAP data viewer (Fig. 4(b)) at marker 6(a) location show the same region and features as the vehicle moves back and forth. The false negative loop closure detection could be attributed to the camera's field of view. In contrast, using LIO-SAM, marker 6(b) demonstrates a successful global loop closure detection (true positive).

Fig. 3(g) shows the top view of A3, while Fig. 3(h) and Fig. 15(i) show the mapping trajectory of A3, for RTABMAP and LIO SAM respectively. From the path in Fig. 3(h) and Fig 3(i), no global close loops occur with RTABMAP and LIO-SAM. Marker 7(a) indicates that there is no global loop

because the last frames do not receive enough inliers to form a closed loop with the starting frame, as shown in Fig. 4(c). Fig.5(a) and Fig. 5(b) provides a close-up view of marker 6(b) before and after applying the LIO-SAM calculation, where the purple line indicates the loop constraint that effectively closes the loop. However, there are some artifacts leading to duplicate features in the same location, resulting in an inaccurate map, as shown in Fig. 5(c). For marker 7(b), Fig. 5(d) shows the close-up of the global map. It also shows that no global loop closure detected based on the output point cloud. The point cloud of the mapping trajectory does not overlap with the original point cloud resulting in inaccurate representation of the map.

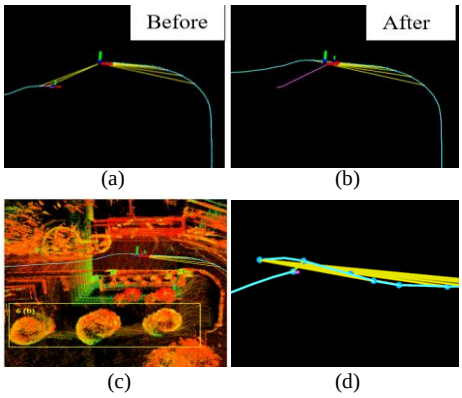


Figure 5(a) Before and (b) after loop close in A3 for marker 6(b) for LIO SAM. (c) LIO-SAM: Artifacts presence in the map (d) LIO - SAM: Close up view from marker 7(b)

IV. Conclusions And Discussion

Both LIO-SAM and RTAB-Map use Lidar scans to create 3D maps, but their approaches to loop closure detection differ. LIO-SAM primarily relies on Lidar scans for odometry and loop closure detection. In contrast, RTAB-Map uses Lidar data for odometry but relies on visual information from camera images for loop closure. RTAB-Map compares visual features to identify loop closures, which can be limited by factors like field of view, lighting variations, and object similarities, leading to potential detection errors. LIO-SAM utilizes raw Lidar scans for odometry, eliminates noise, and extracts key points for loop closure using the Iterative Closest Point algorithm. It employs a radius-search method to align scans from different parts of the trajectory, correcting for drift for loop closure detection. LIO-SAM employs a tightly coupled approach with the IMU, achieving better loop closure detection by fusing Lidar and IMU data to match scans and correct drift. However, tightly coupling can introduce errors if either sensor has inaccuracies. Even with loop closure, map artifacts can still occur due to false positives, false negatives, sensor quality, environmental factors, and handling of dynamic changes.

Acknowledgment

The authors thank the MIMOS Berhad for providing facilities and all colleagues at Advance Intelligence Lab for contributing in this research work. This research is funded by Malaysian Ministry of Science, Technology and Innovation (MOSTI) under the strategic research grant.

References

[1] A. Ismail, "The Effect of Labour Shortage in the Supply and Demand of Palm Oil in Malaysia," Oil Palm Industry Economic Journal, vol. 13, no. 2, pp. 1-26, 2013.

[2] R. Radmanesh, Z. Wang, V. S. Chipade, G. Tsechpenakis, and D. Panagou, "LIV-LAM: LiDAR and visual localization and mapping," in 2020 American Control Conference (ACC), 2020: IEEE, pp. 659-664.

[3] G. Bresson, Z. Alsayed, L. Yu, and S. Glaser, "Simultaneous localization and mapping: A survey of current trends in autonomous driving," IEEE Transactions on Intelligent Vehicles, vol. 2, no. 3, pp. 194-220, 2017.

[4] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," IEEE Transactions on Robotics, vol. 24, no. 6, pp. 1365-1378, 2008.

[5] F. Dellaert, "Factor graphs and GTSAM: A hands-on introduction," Georgia Institute of Technology, 2012.

[6] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g 2 o: A general framework for graph optimization," in 2011 IEEE International Conference on Robotics and Automation, 2011: IEEE, pp. 3607-3613.

[7] M. Labbe and F. Michaud, "Online global loop closure detection for large-scale multi-session graph-based SLAM," in 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2014: IEEE, pp. 2661-2666.

[8] S. Patra, H. Aggarwal, H. Arora, S. Banerjee, and C. Arora, "Computing egomotion with local loop closures for egocentric videos," in 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), 2017: IEEE, pp. 454-463.

[9] M. Labbe, "Simultaneous Localization and Mapping (SLAM) with RTAB-Map," Département de Génie Électrique et Génie Informatique, 2015.

[10] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. Montiel, and J. D. Tardós, "ORB-SLAM3: An accurate open-source library for visual," Visual-Inertial and Multi-Map SLAM, 2020.

[11] 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), 2018: IEEE, pp. 4758-4765.

[12] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in 2020 IEEE/RSJ international conference on intelligent robots and systems (IROS), 2020: IEEE, pp. 5135-5142.

[13] H. T. Warku, N. Y. Ko, H. G. Yeom, and W. Choi, "Three-Dimensional Mapping of Indoor and Outdoor Environment Using LIO-SAM," in 2021 21st International Conference on Control, Automation and Systems (ICCAS), 2021: IEEE, pp. 1455-1458.

[14] T. Shan, B. Englot, C. Ratti, and D. Rus, "Lvi-sam: Tightly-coupled lidar-visual-inertial odometry via smoothing and mapping," in 2021 IEEE international conference on robotics and automation (ICRA), 2021: IEEE, pp. 5692-5698.

[15] K. J. de Jesus, H. J. Kobs, A. R. Cukla, M. A. d. S. L. Cuadros, and D. F. T. Gamarra, "Comparison of visual SLAM algorithms ORB-SLAM2, RTAB-Map and SPTAM in internal and external environments with ROS," in 2021 Latin American Robotics Symposium (LARS), 2021 Brazilian Symposium on Robotics (SBR), and 2021 Workshop on Robotics in Education (WRE), 2021: IEEE, pp. 216-221.

[16] M. Kim, M. Zhou, S. Lee, and H. Lee, "Development of an Autonomous Mobile Robot in the Outdoor Environments with a Comparative Survey of LiDAR SLAM," in 2022 22nd International Conference on Control, Automation and Systems (ICCAS), 2022: IEEE, pp. 1990-1995.

[17] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the Bayes tree," The International Journal of Robotics Research, vol. 31, no. 2, pp. 216-235, 2012.

[18] G. Wang, X. Gao, T. Zhang, Q. Xu, and W. Zhou, "LiDAR Odometry and Mapping Based on Neighborhood Information Constraints for Rugged Terrain," Remote Sensing, vol. 14, no. 20, p. 5229, 2022