2021 International Conference on Signal Processing and Machine Learning (CONF-SPML) | 978-1-6654-1734-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/CONF-SPML54095.2021.00040

Review on LiDAR-based SLAM Techniques

Leyao Huang
Logistics Engineering College
Shanghai Maritime University
Shanghai, China
Hleyao0204@outlook.com

***Abstract*—LiDAR-based Simultaneous Localization and Mapping (LiDAR-SLAM) uses the LiDAR sensor to localize itself by observing environmental features and incrementally build the map of the surrounding environment. In this way, the purpose of simultaneous localization and mapping in the unknown environment can be achieved. Localization and mapping with high robustness, high accuracy, and high practicability is a complex and hot issue in recent years. This paper will briefly introduce the information background, classification and development history of LiDAR-SLAM. We will also summarize the common frameworks of LiDAR-SLAM and the function of core modules in the existing LiDAR-SLAM. Additionally, the state-of-the-art multi-sensor fusion-based LiDAR-SLAM techniques are investigated, and the future development trend of LiDAR-SLAM is discussed.**

***Keywords*—LiDAR-SLAM, LiDAR Odometry, LiDAR Mapping, Multi-sensor fusion, Point Cloud**

I. Introduction

The past decades have witnessed the rapid development of autonomous navigation of mobile robots. In order to enable the robot to better localize itself and build the structure consistency map of the environment without a priori environment information, SLAM technology enables robots to explore unknown environments and is the basis for autonomous decision-making, planning and control of robots. With SLAM technology, autonomous mobile robots can be used to explore areas on the ground, such as terrain, the ocean floor, air space and places where humans may not be able to reach or are potentially dangerous to humans.

At present, the main SLAM techniques include LiDAR-SLAM and Visual SLAM (V-SLAM). V-SLAM takes the camera as the main sensor and the image as the main information source. It has many advantages such as rich information, low cost, light weight and small size. However, due to the camera's sensitivity to light and perspective, as well as the relatively difficult ability to extract 3D information, its disadvantage is that the effect of V-SLAM is not as ideal as that of LiDAR-SLAM in the case of unstable ambient lighting and light texture, such as outdoor environment with unstable light. To solve these problems, the recent achievements in this field include illumination model modification and high robustness feature learning based on deep learning algorithms [1] [2]. V-SLAM has achieved relatively ideal results in some indoor circumstances at present, and with the progress of these technologies, it is predicted to have a more stable performance in the environment with changing illumination in the future.

Compared with the V-SLAM, LiDAR-SLAM can accomplish its own localization and mapping under more diversified environmental conditions. Object information collected by lidar presents a series of scattered points with accurate angle and distance information, which is called a point cloud. In general, the LiDAR-SLAM system extracts and matches the features of two consecutive point-cloud graphs of different moments, calculate the relative motion distance and pose change through the LiDAR odometry, and thus complete the positioning of the robot itself.

The significant advantages of the LiDAR sensor are the relatively accurate range measurement, simple error model, stable operation in the environment outside the direct glare of strong light, and relatively easy point cloud processing. At the same time, the point cloud information contains direct geometric relations, which makes the path planning and navigation of the robot intuitive. Recent developments in sensing hardware have reduced the LiDAR size and weight. The LiDAR, which can be carried in a backpack or even mounted on a tiny aircraft, is more commonly used in autonomous navigation.

There are many categories of LiDAR-SLAM. Firstly, since different types of LiDAR sensors are used, LiDAR-SLAM can be divided into 2D Lidar-SLAM and 3D Lidar-SLAM. 2D LiDAR has a lower cost and is suitable for a simple indoor environment. 3D LiDAR has a high cost, is suitable for outdoor surveying and mapping, and has a large amount of information. The differences between specific frameworks and sub-modules will be explained in detail in Section 2 of this paper.

As for the back end, the first mainstream approach is based on filter, similar to the Bayesian estimation process. Among them, the representative algorithm is Gmapping [3], which only estimates the position of the robot at the current moment, so the calculation is small. This is also the disadvantage of this method because once the error is generated at the previous moment, the error cannot be repaired. When the environment is large, the effect of mapping in this method may be non-ideal to build a map. The second method is mainly based on non-linear optimization, among which, graph optimization is widely used in LiDAR-SLAM. In graph-based SLAM, the robot's pose is a node or vertex, and the relationship between the poses constitutes the edge. The errors accumulated in the map building process are optimized by the nonlinear least square method to obtain the final map. One of typical algorithms is Cartographer [4].

Fast SLAM [5], proposed in 2002, is the earliest system capable of a real-time output of raster maps, which has the problems of large memory consumption and serious particle dissipation. To solve this defect, Optimal RBPF [6] was proposed to optimize the Gmapping algorithm and further reduce the particle degradation problem. As for the graph-based optimization scheme, Karto-slam [7] was the initial system which introduced the concept of sparsity and laid the foundation of Cartographer. Basing on the emergence of graph optimization method, Cartographer builds on its improved accuracy and robustness and reduces the amount of computation, therefore several early filter-based methods have been instead.

With the publication by Zhang J et al of LOAM [8], a pure lidar algorithm based on 3D LiDAR, the foundation of 3D Lidar-SLAM has been laid. On this basis, V-LOAM [9] integrates the vision lidar, which improves the accuracy and

978-1-6654-1734-1/21/$31.00 ©2021 IEEE
DOI 10.1109/CONF-SPML54095.2021.00040

robustness of the algorithm but still does not consider the loop-closure detection. Subsequently, it is LIO-SAM [10] that uses inertial odometry, which integrates IMU and LOAM optimization frameworks. In recent years, some newer algorithms like IMLS and LVIO have integrated other sensors to solve many of the degradation problems in LOAM.

So far, 2D Lidar-SLAM has developed relatively well, while 3D Lidar-SLAM still has many problems to be solved and improved.

II. Frameworks for LiDAR-SLAM Systems

A. System Structure of the 2D LiDAR-SLAM system

A completed 2D LiDAR-SLAM involved the front-end incremental estimation part and the back-end mapping part. In regard to the front-end, it consists of two products, obtaining the data of the sensor and scan matching. For the back-end, the main task is to solve a nonlinear optimization problem. Common 2D SLAM methods include EKF SLAM, [3], [4], [5], [11],[12] and so on.

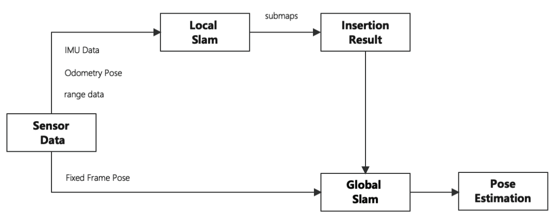

Fig. 1. Framework of the Cartograph

Take a current mainstream model of 2D LiDAR-SLAM, Cartographer, as an example to explain how 2D LiDAR-SLAM works. Cartographer launched by Google as a set of graph-based optimization SLAM algorithms provides a solution for real-time indoor map construction. It proposed a new loop-closure detection method based on LiDAR data, which can reduce the calculation, meet the need of large space map building, and optimize the large-scale data in real-time. The main parts of Cartograph mainly consist of local slam and global slam.

As is shown in figure 1, local SLAM produce utilizes odometry and IMU data to calculate the trajectory and give the estimated value of the robot pose, which is used as the initial value to match the LiDAR data and update the pose estimator. Then, frames of lidar are motion-filtered and stacked to form submaps. In the global SLAM part, loop-closure detection and back-end optimization are finished to form a complete usable map of the world map. In this way, the pose of all scans and submaps can be optimized, following sparse attitude adjustment.

B. System Structure of the 3D LiDAR-SLAM system

In the past several decades, from EKF slam to Cartographer, 2D Slam is well developed. However, some challenging environments, such as smoke and drops of water in the air still have a great impact on the precision of effect. With the introduction of LOAM by Zhang J et al.[8], 3D SLAM addresses the influence of many environmental factors. 3D SLAM usually takes, 3D LiDAR data and odometer data as input data, and 3D point cloud map, robot trajectory or pose graph as output.

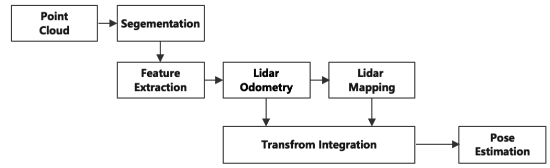

Fig. 2. Framework of the Lego-LOAM

In this paper, Lego-LOAM will be introduced to explain the common framework of 3D LiDAR-SLAM. As is shown in figure 2, the whole system can be divided into 5 parts.

The segmentation module removes the ground points in the receiving point cloud and synthesizes the remaining point cloud to an image. Then, based on the method of image segmentation, the distance image is separated into multiple clusters and labels are assigned to remove the interference of noise points. Next, feature extraction in Lego-LOAM is similar to LOAM, which is based on smoothness to sort the points of the segmented image to determine the plane points and corner points. This method can greatly reduce the amount of calculation in the subsequent process. The difference is that Lego-LOAM adds ground points based on smoothness. The subsequent module, LiDAR odometry, estimates the motion of the sensor/robot between two consecutive scans and uses Levenberg-Marquardt (LM) methods two times to obtain the transformation information of all corresponding points. In the Lidar mapping module, Lego-LOAM registers the feature points with the surrounding point cloud image and carries out further attitude changes at a low frequency, which is similar to LOAM. By contrast, Lego-loam chooses to store each set of feature points instead of a single point cloud. In addition to the basic module, a major improvement in Lego-Loam is the addition of loop-closure detection to reduce errors and improve SLAM accuracy, while it also increases the calculation effort of the system.

C. The core problems of LiDAR-SLAM system

Due to the disordered, sparse and limited information characteristics of a point cloud, there is still a lot of room for the improvement of the existing LiDAR-SLAM system. In the field of LiDAR-SLAM, 3D SLAM is objectively a more challenging algorithm than 2D SLAM, because the amount of data and computation is much larger, the feature matching in spatial dimension is more complex than that in a plane, and the demand for positioning accuracy is also higher. However, 3D SLAM has indispensable research significance at present. First, by means of 3D point positions, it can solve the problem of losing points in the non-localization environment. Second, when a robot is in complex terrains, 3D SLAM can perform relatively better.

1) Point cloud pre-processing. Point cloud acquisition has the advantages of fast, strong penetration, strong real-time, strong dynamic, high density and high efficiency, so its development prospects and needs are very broad. When get point cloud data model in practice, because of the complicated terrain object shade, light problems, impression, hard to avoid can gather the noise of the unreasonable points, it is difficult to complete measurements meanwhile makes difficult to meet the requirements of the model precision and real-time, so that the point cloud pretreatment process has a great significance.

In the module of Feature extraction, the basic process is to divide the depth map into equal submaps in the horizontal direction. Then, through the smoothness formula, the smoothness data (c) is obtained and sorted, and the edge points ($c > c_{th}$) and flat points ($c < c_{th}$) are differentiated according to the threshold value. Sort the points in the set and find the smallest point as the plane point and the largest point as the edge point. After screening, the refined edge point F_e and the flat point set F_p were obtained.

In addition to the curvature method used in LOAM system, feature vectors, histograms, rotated images and other good methods, but there are large computational problems.

2) Lidar Odometry. [2] use the label information obtained by Segmentation to find the associated points for label matching, and then perform a two-step LM optimization. Compared with the Loam method which only carries out one LM optimization, the accuracy is improved. To simplify the algorithm of LOAM, A-loam removes the IMU part of LOAM in the LiDAR odometry module and uses Eigen for bit conversion and uses Ceres for iterative optimization. With 3D depth sensors, Deschaud [13] developed the Implicit Moving Least Squares (IMLS), a low-drift SLAM similar to LOAM, which was based only on 3D LiDAR data. Histogram used to be a common method for visual odometry while a histogram of point heights is also can be used for describing locations in LiDAR-SLAM. In addition, there is also a method to create descriptors and calculate density functions by dividing point clouds into a solid grid [14], which describes the shape and is later merged into the global descriptors in the histogram, and the map will be represented as a set of points along the trajectory, and each point has a defined origin to build a similar graph structure. However, when the robot is between or outside the vertices, the exact position is not easy to retrieve. Apart from the method of the histogram, local descriptors may also be used for scene recognition. Bosse and Zlot [15] use descriptors calculated around randomly selected key points in the global map and local frames. In a similar approach, Tombari and Di Stefano[16] did not divide the global point cloud into a set of locations but instead looked for all the corresponding ones instead. These correspond to instances that are then clustered as locations using Hough voting. Dube et al. [17] recently proposed that recognition sites are extracted and matched based on 3D segmentation, providing a balance between local and global descriptors. [18] It has been demonstrated that this strategy can be applied to real-time multi-robot SLAM applications. With the development of deep learning, DMLO[19] algorithm uses global sparse matching to realize LiDAR odometry.

3) Lidar Mapping. The interframe matching algorithm directly affects the LiDAR SLAM, and the back-end optimization only eliminates the errors accumulated during the process. The more accurate the interframe matching estimation is, the better the image construction effect is.

ICP(Iterative Closest Point), NDT(Normal Distribution Transformation), PI-ICP(Point-to-Line Iterative Closest Point), CSM(Correlation Scan Match) is commonly used at present.

4) Loop-closure detection and global optimization

One of the great challenges in loop-closure detection is that images of the same location can be highly different because of the factors like illumination. Finding a representative feature to distinguish the identification points is the key problem.

Before the era of deep learning, the traditional approach is to treat the recognition problem as an image retrieval task, use invariant features to represent each image, and then converge the features into vectors, such as a bag of words (BoW), Vector of locally aggregated Descriptors (VLAD), Fisher Vector (FV), etc.

The core idea of BoW method is to use clustering method to train a sample after extracting keypoint descriptors. Then the number of times that each descriptor vector in each image appears in the sample is used to represent the image. The disadvantage of this method is that it requires a large sample size. The core idea of FV method is to use Gaussian mixture model (GMM) to represent each image by calculating the mean value, covariance and other parameters in the GMM. The advantage of this method is a high accuracy, but the disadvantage is a large amount of calculation.

VLAD, introduced by Jegou et al., the central idea is aggregated, a simplified form of FV. The main method is to train a small codebook by clustering method, find the nearest cluster center of the codebook for each feature in the image, and then add up the difference between all features and the cluster center.

These traditional methods are not navigable, so they cannot carry out back propagation. With the popularization of deep learning, they are gradually replaced by CNN.

Based on VLAD algorithm, an improved method is proposed in the All About VLAD . Then, NetVLAD [20]is proposed based on deep convolutional neural networks.

NetVLAD proposes a convolutional neural network structure that can realize end-to-end recognition. The main innovation point is to embed the traditional VLAD structure into the CNN network structure to get a new VLAD layer[21], which can be optimized using back propagation.

For LiDAR-SLAM, pointNetVLAD[22] has designed a new neural network to solve the irregular point cloud problem. It is a combination/modification of existing PointNet and NetVLAD, which allows end-to-end training and inference to extract global descriptors from a given 3D point cloud.

LDP-Net [23] has a significant performance improvement over PointNetVLAD. This network extracts discriminant and generalized global descriptors from original 3D point clouds and applies them to point cloud-based retrieval tasks to realize large-scale location recognition.

III. Multi-Sensor-ided 3D LiDAR-SLAM Approaches

With the increasing demand for SLAM, the single-sensor SLAM system is limited in many aspects. For the sensor, lidar is typical of low vertical resolution and sparse point clouds making it tricky for feature tracking in the environment with few textures or erratic weather and lighting. Second, if the LiDAR on the mobile robot is disturbed by the fierce movement, the sensing accuracy will be directly affected. In the practical application, some cases of LiDAR degradation cannot be ignored. Because of the few sensing points, matching LiDAR features can easily lead to unconstrained attitude problems. Besides, some previously useful SLAM approach is often rendered useless, as the low frequency of LiDAR updates limits the task applications that need to respond quickly. In terms of positioning, some existing LiDAR-SLAM systems are difficult to build map and positioning at the same time. Moreover, due to the complexity of the environment, there is a certain complexity in the practical application of Landmark.

For these challenges brought by the defects of lidar sensor and the environment to SLAM, it is difficult for a pure-sensor SLAM to meet our needs for speediness and accuracy, so it is necessary to introduce multi-sensor-aided SLAM.

A. Inertial-aided LiDAR-SLAM

The concept of IMU was first proposed in [24], combining many inertial measurements between two key frames into a single relative motion constraint. [25] propose a pre-integration theory to provide a more formal form of rotation noise, avoiding the start of a row of rotatable representations, and applying IMU pre-integration theory to factor graph models makes incremental smoothing algorithms more possible.

Joint pose estimation of lidar and IMU sensors has been widely studied and can be divided into two types. The first is loosely coupled fusion, which is characterized by considering lidar and IMU estimates separately. For example, in classical LOAM [8], the main function is to remove the motion distortion of Lidar point clouds. Specifically, the algorithm assumes that the IMU velocity is zero to assist the IMU calculation and the LiDAR odometry and then decouples the LiDAR and IMU measurements, mainly taking the IMU as a prior for the entire system. In [26], a Loosely Coupled extended Kalman filter (EKF) is used to fuse IMU and 2D LiDAR. Extending this to 3D LiDAR, Lynen et al. [27] provided a modular approach to the fusion of IMU measurements with other relative pose measurements, such as from a camera, Lidar or even pressure sensors, via EKF in 3D situations.

Loosely-Coupled method is computationally efficient, but its accuracy is much lower than that of tight-coupled method [28]. For 2D plane motion estimation, [29] the method is to extract and match the lines in the 2D laser scan, and a Kalman filter is used to correct the IMU state in the LiDAR measurement domain. Hemann[30] et al. proposed a method for tightly coupled IMU propagation and accumulation of lidar height maps in the form of error state Kalman filters. The updated status is modified by matching the lidar elevation map and the prior digital elevation model (DEM). This method can make the robot work without prior map information because it has no GPS navigation capability over a long distance. This method has also been used in R-Lins [31]. In [32][33], the original measurements directly from the IMU and the predicted IMU measurements from the continuous trajectory are used to calculate the remaining amount to be optimized.

LIO-mapping [34] is a classic LIO system that provides more reliable estimates by jointly minimizing the cost from LiDAR and IMU measurements. The LIO can be performed well and with less error in long-term experiments, even avoiding the problem of lidar measurement degradation in complex cases with fast motion conditions or inadequate features. In addition, the rotation constraint optimization algorithm is used to further align the lidar pose with the global map to obtain a more reliable pose estimation.

LIO-sam [35] added IMU pre-integration factor and GPS factor on the basis of Lego-Loam, removed the frame-matching part and refined the design of frame matching part to obtain the globally consistent pose of the robot. Experimental results on different data sets show that the tight coupling of various sensors improves the robustness of the motion estimation.

LIO can not rely on prior maps, so the accuracy does not change with the change of the environment, but the accuracy is low at the same time. In order to avoid the problem of low precision of LIO, in [36], based on Cartographer, LiDAR global module which is sensitive to environmental changes but with high accuracy is used to complement each other, greatly improving the accuracy and solving the problem of accurate positioning in dynamic changing scenes.

B. Visual-aided LiDAR-SLAM

In nearly two years of research, a LIO system can help to correct the point cloud distortion and solve the problem of feature loss in a short time. However, pure LiDAR- SLAM is still difficult to meet some high information demand environments. To further improve system performance, the fusion of LiDAR, camera, and IMU measurements is attracting increasing attention. In the first section, it is mentioned that the vision-based method is more suitable for scene recognition and better in texture recognition, while the LiDAR can provide more accurate distance information. By combining these two kinds of data, considerable data can be obtained due to their complementary characteristics. At the same time, the measurement scale and attitude can be recovered by the IMU measurement to assist the vision-inertial system.

The current research direction can be summarized as achieving high precision in filtering algorithms and pursuing real-time the optimization algorithm. Similar to LIO, when IMU is added, the research direction is also divided into loose coupling and tight coupling. Loose coupling is similar to filtering fusion, and the state measured by IMU and the state obtained by visual odometry are calculated separately and then fused, while tight coupling is based on nonlinear optimization. IMU measurement and visual constraint information are optimized in a nonlinear optimization function. The tightly coupled framework allows the IMU data to correct the visual odometry, and the visual odometry information can also correct the IMU's zero offsets, so the tightly coupled framework provides a higher accuracy.

The SLAM system of SOTA using only LiDAR scanner is [37]-[40], which often requires a motion model. The method in [41] combines a binocular camera and a LiDAR scanner. It has motion estimation from VO and is refined by matching LiDAR frames. Meanwhile, VIO is more robust and accurate than VO[42]. VLOAM[9] uses IMU motion prediction and visual-inertial coupling motion estimation, and then a scanning matching method further revises the motion estimation and registers the map. The result is a high frequency, low delay motion estimation and a dense, accurate 3D map registration. In addition, the method can solve the problem of single sensor failure by skipping the failed module. Similar to [43] in terms of system composition, [43] incorporates the ideas of [44] and [25]. Based on LOAM, tightly coupled binocular VIO based on IMU pre-integration and unstructured visual factor is used. The VIO output of IMU frequency is used for point cloud distortion removal. In addition, LiDAR's loop-closure can be enhanced. One difference [43] is that a tightly coupled VIO is used as the motion model to initialize the LiDAR mapping algorithm while VLOAM uses a loosely coupled IMU and camera.

In addition to the multi-sensor fusion based on VLOAM, other ideas have been proposed.

LIC-fusion [45] proposed the optimal (maximum linear error) multi-mode sensor Fusion scheme, which effectively fused the edge/plane feature points detected by LiDAR into the VIO framework based on MSCKF. LIMO [46] LiDAR is used to enhance the depth of visual features by fitting the local plane and performs well in autonomous driving scenes. DEMO [47] uses sparse LiDAR point cloud/depth images and triangulation to obtain the depth of feature points, respectively. LVI-SAM [48] proposes a tightly coupled LiDAR-vision-inertial odometry through smoothing and mapping, which realizes the global optimization of scene recognition assistance. The LVI-SAM consists of a vision-inertial system (VIS) and a LiDAR- inertial system (LIS). The two subsystems are designed to be tightly coupled, with the former using the latter estimates to facilitate initialization and the latter using the former estimates to make initial guesses to support scan-matching. The framework is robust to sensor degradation by bypassing the failed subsystem through fault detection.

From the perspective of current development, pure vision or pure LiDAR SLAM has been increasingly involved in large-scale positioning, navigation and three-dimensional or semantic map construction, ground robots, drones, VR/AR/MR, AGV and other needs. Multi-sensor fusion is definitely the development trend of SLAM in the future. Loose coupling algorithm performs better in real time, while tight coupling algorithm is reflected in accuracy. In different combinations, data modeling, state analysis and optimization between multiple sensors are still worth exploring in order to meet different requirements. For example, from the industrial point of view, the discrete system has always been a big difficulty. If LiDAR-SLAM is to be put into use, multi-machine collaboration is inevitable.

IV. Conclusion

LiDAR-SLAM is a challenging subject that has made amazing progress over the past few decades. Currently, 2D SLAM has been relatively complete. By contrast, although some classic frameworks of 3D SLAM can solve scene recognition and map building in some simple environments, there are still many development spaces and problems to be solved.

At present, closed-closure detection is still a core problem of LiDAR-SLAM. When the errors at the front-end accumulate to a certain extent, global consistency and feature extraction will become difficult, and the computational cost of ICP is also large. In terms of feature extraction, unlike visual SLAM, which has a library to call, LiDAR-SLAM does not have a relatively uniform standard method. In terms of hardware, LiDAR-SLAM is asked to be popularized, the cost should be taken into account. The price of the better laser sensor is still relatively expensive, whereas the low-cost sensor may affect the performance of the test. Reducing the cost of LiDAR while maintaining performance is also a challenge.

In recent years, the main research content is mainly focused on multi-sensor-aided methods, which can definitely be the focus of the future development of LiDAR-SLAM. Currently, visual-aided and inertial-aided LiDAR-SLAM has been widely recognized. Among them, the major challenges focus on the fusion and synchronization of information from different sensors and how to achieve lightweight and fast-response SLAM after the introduction of multiple sensors. With the rapid development of deep learning, it is also an approach to multi-source fusion SLAM by extracting semantic information and fusing geometric features, through inputting images and point cloud information into neural networks. However, this idea is still in its embryonic stage, and there is still plenty of space for expansion.

In short, 3D LiDAR-SLAM is developing towards directions of higher accuracy, higher real-time performance, and high generalization suitable for different environments.

References

[1] C. Park and J. Song, "Illumination change compensation and extraction of corner feature orientation for upward-looking camera-based SLAM," 2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Goyangi, Korea (South), 2015, pp. 224-227.

[2] S. Lee, H. Jo, H. M. Cho and E. Kim, "Visual Loop Closure Detection over Illumination Change," 2019 16th International Conference on Ubiquitous Robots (UR), Jeju, Korea (South), 2019, pp. 77-80.

[3] G. Grisettiyz, C. Stachniss and W. Burgard, "Improving Grid-based SLAM with Rao-Blackwellized Particle Filters by Adaptive Proposals and Selective Resampling," Proceedings of the 2005 IEEE International Conference on Robotics and Automation, Barcelona, Spain, 2005, pp. 2432-2437.

[4] W. Hess, D. Kohler, H. Rapp and D. Andor, "Real-time loop closure in 2D LIDAR SLAM," 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 2016, pp. 1271-1278.

[5] W. Shiguang and W. Chengdong, "An improved FastSLAM2.0 algorithm using Kullback-Leibler Divergence," 2017 4th International Conference on Systems and Informatics (ICSAI), Hangzhou, China, 2017, pp. 225-228.

[6] J. Sheng, S. Jia, S. Tano and X. Li, "RBPF-based map generation for an indoor mobile robot," SICE Annual Conference 2011, Tokyo, Japan, 2011, pp. 1666-1669.

[7] K. Konolige, G. Grisetti, R. Kümmerle, W. Burgard, B. Limketkai and R. Vincent, "Efficient Sparse Pose Adjustment for 2D mapping," 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 2010, pp. 22-29.

[8] J. Zhang and S. Singh, "LOAM: Lidar Odometry and Mapping in Real-time", Proceedings of Robotics: Science and Systems, 2014.

[9] J. Zhang and S. Singh, "Visual-lidar odometry and mapping: low-drift, robust, and fast," 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 2015, pp. 2174-2181.

[10] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti and D. Rus, "LIO-SAM: Tightly-coupled Lidar Inertial Odometry via Smoothing and Mapping," 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 2020, pp. 5135-5142.

[11] R. Smith, M. Self and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," Proceedings. 1987 IEEE International Conference on Robotics and Automation, Raleigh, NC, USA, 1987, pp. 850-850.

[12] T. Röhling, J. Mack and D. Schulz, "A fast histogram-based similarity measure for detecting loop closures in 3-D LIDAR data," 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 2015, pp. 736-741.

[13] J. Deschaud, "IMLS-SLAM: Scan-to-Model Matching Based on 3D Data," 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, 2018, pp. 2480-2485.

[14] M. Magnusson, H. Andreasson, A. Nuchter, and A. J. Lilienthal, "Appearance-based loop detection from 3d laser data using the normal distributions transform," in Robotics and Automation, 2009. ICRA'09. IEEE International Conference on. IEEE, 2009, pp. 23–28.z

[15] M. Bosse and R. Zlot, "Place recognition using keypoint voting in large 3d lidar datasets," in Robotics and Automation (ICRA), 2013 IEEE International Conference on. IEEE, 2013, pp. 2677–2684.

[16] F. Tombari and L. Di Stefano, "Object recognition in 3d scenes with occlusions and clutter by hough voting," in Image and Video Technology (PSIVT), 2010 Fourth Pacific-Rim Symposium on. IEEE, 2010, pp. 349–355.

[17] R. Dube´, D. Dugas, E. Stumm, J. Nieto, R. Siegwart, and C. Cadena, "Segmatch: Segment Based Place Recognition in 3D Point Clouds," IEEE International Conference on Robotics and Automation, 2017, pp. 5266–5272.

[18] Cadena, "An Online Multi-Robot SLAM System for 3D LiDARs," IEEE International Conference on Intelligent Robots and Systems (IROS), 2017

[19] Z. Li and N. Wang, "DMLO: Deep Matching LiDAR Odometry," 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 2020, pp. 6010-6017.

[20] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla and J. Sivic, "NetVLAD: CNN Architecture for Weakly Supervised Place Recognition," 2016

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 5297-5307.

[21] R. Arandjelović, P. Gronat, A. Torii, T. Pajdla and J. Sivic, "NetVLAD: CNN Architecture for Weakly Supervised Place Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 6, pp. 1437-1451, 1 June 2018.

[22] M. A. Uy and G. H. Lee, "PointNetVLAD: Deep Point Cloud Based Retrieval for Large-Scale Place Recognition," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 4470-4479.

[23] Z. Liu et al., "LPD-Net: 3D Point Cloud Learning for Large-Scale Place Recognition and Environment Analysis," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 2831-2840

[24] T. Lupton and S. Sukkarieh, “visual–inertial-aided navigation for high-dynamic motion in built environments without initial conditions,” IEEE Trans. Robot., vol. 28, no. 1, pp. 61–76, Feb. 2012

[25] C. Forster, L. Carlone, F. Dellaert, D. Scaramuzza. On-Manifold Preintegration for Real-Time Visual-Inertial Odometry. IEEE Transactions on Robotics, 33(1), 1-21, 2017.

[26] J. Tang, Y. Chen, X. Niu, L. Wang, L. Chen, J. Liu, C. Shi, and J. Hyyppa ̈, “Lidar scan matching aided inertial navigation system in gnss-denied environments,” Sensors, vol. 15, no. 7, pp. 16 710–16 728, 2015.

[27] S. Lynen, M. W. Achtelik, S. Weiss, M. Chli, and R. Siegwart, “A robust and modular multi-sensor fusion approach applied to mav navigation,” in Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on. IEEE, 2013, pp. 3923–3929.

[28] M. Li, B. H. Kim, and A. I. Mourikis, “Real-time motion tracking on a cellphone using inertial sensing and a rolling-shutter camera,” in Robotics and Automation (ICRA), 2013 IEEE International Conference on. IEEE, 2013, pp. 4712–4719.

[29] A. Soloviev, D. Bates, and F. Van Graas, “Tight coupling of laser scanner and inertial measurements for a fully autonomous relative navigation solution,” Navigation, vol. 54, no. 3, pp. 189–205, 2007.

[30] [35] G. Hemann, S. Singh, and M. Kaess, “Long-range gps-denied aerial inertial navigation with lidar localization,” in Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on. IEEE, 2016, pp. 1659–1666.

[31] C.Park,P.Moghadam,S.Kim,A.Elfes,C.Fookes,andS.Sridharan, “Elastic lidar fusion: Dense map-centric continuous-time slam,” in 2018 IEEE International Conference on Robotics and Automation(ICRA). IEEE, 2018, pp. 1206–1213.

[32] M. Bosse and R. Zlot, “Continuous 3d scan-matching with a spinning 2d laser,” in Robotics and Automation, 2009. ICRA’09. IEEE International Conference on. IEEE, 2009, pp. 4312–4319.

[33] C.Park,P.Moghadam,S.Kim,A.Elfes,C.Fookes,andS.Sridharan, “Elastic lidar fusion: Dense map-centric continuous-time slam,” in 2018 IEEE International Conference on Robotics and Automation

[34] H. Ye, Y. Chen, M. Liu. Tightly Coupled 3D LiDAR Inertial Odometry and Mapping. IEEE International Conference on Robotics and Automation, 3144-3150, 2019.

[35] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, D. Rus. LIO-SAM: Tightly-coupled Lidar Inertial Odometry via Smoothing and Mapping, IEEE/RSJ International Conference on Intelligent Robots and Systems, 2020.

[36] W. Ding, S. Hou, H. Gao, G. Wan, S. Song. LiDAR Inertial Odometry Aided Robust LiDAR Localization System in Changing City Scenes. IEEE International Conference on Robotics and Automation, 4322-4328, 2020.

[37] S. Ceriani, C. Snchez, P. Taddei, E. Wolfart, and V. Sequeira, “Pose interpolation slam for large maps using moving 3d sensors,” in 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 750–757, Sept 2015.

[38] M. Velas, M. Spanel, and A. Herout, “Collar line segments for fast odometry estimation from velodyne point clouds,” in 2016 IEEE International Conference on Robotics and Automation (ICRA), pp. 4486–4495, May 2016.

[39] C. Wei, T. Wu, and H. Fu, “Plain-to-plain scan registration based on geometric distributions of points,” in 2015 IEEE International Conference on Information and Automation, pp. 1194–1199, Aug 2015.

[40] J.-E. Deschaud, “Imls-slam: scan-to-model matching based on 3d data,” CoRR, vol. abs/1802.08633, 2018.

[41] D. Droeschel, J. Stckler, and S. Behnke, “Local multi-resolution representation for 6d motion estimation and mapping with a continuously rotating 3d laser scanner,” in 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 5221–5226, May 2014.

[42] J. A. Delmerico and D. Scaramuzza, “A benchmark comparison of monocular visual-inertial odometry algorithms for flying robots,” 2018.

[43] W. Shao, S. Vijayarangan, C. Li, G. Kantor. Stereo Visual Inertial LiDAR Simultaneous Localization and Mapping. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2019.

[44] J. Zhang and S. Singh, “Laser-visual-inertial odometry and mapping with high robustness and low drift,” 08 2018.

[45] X. Zuo, Y. Yang, P. Geneva, J. Lv, Y. Liu, G. Huang, M. Pollefeys. LIC-Fusion2.0: LiDAR-Inertial-Camera Odometry with Sliding-Window Plane-Feature Tracking.

[46] J. Graeter, A. Wilczynski, M. Lauer. LIMO: LiDAR-Monocular Visual Odometry. IEEE/RSJ International Conference on Intelligent Robots and Systems, 7872-7879, 2018.

[47] J. Zhang, M. Kaess, S. Singh. Real-time Depth Enhanced Monocular Odometry. IEEE/RSJ International Conference on Intelligent Robots and Systems, 4973-4980, 2014.

[48] D. Wisth, M. Camurri, S. Das and M. Fallon, "Unified Multi-Modal Landmark Tracking for Tightly Coupled Lidar-Visual-Inertial Odometry," in IEEE Robotics and Automation Letters, vol. 6, no. 2, pp. 1004-1011, April 2021.