3D LiDAR SLAM : A survey

Article *in* The Photogrammetric Record · May 2024

DOI: 10.1111/phor.12497

CITATIONS
5

READS
1,793

3 authors, including:

Pengcheng Shi
Wuhan University
19 PUBLICATIONS **203** CITATIONS
SEE PROFILE

Jiayuan Li
Wuhan University
64 PUBLICATIONS **1,672** CITATIONS
SEE PROFILE

DOI: 10.1111/phor.12497

REVIEW ARTICLE

3D LiDAR SLAM: A survey

Yongjun Zhang | **Pengcheng Shi** | **Jiayuan Li**

Wuhan University, Wuhan, China

Correspondence
Pengcheng Shi and Jiayuan Li, Wuhan University, Wuhan, China.
Email: shipc_2021@whu.edu.cn and ljy_whu_2012@whu.edu.cn

Funding information
National Key R&D Program of China; Science and Technology Major Project of Hubei Province; National Natural Science Foundation of China; Wuhan university-Huawei Geoinformatics Innovation Laboratory Open Fund, Grant/Award Number: TC20210901025-2023-06

Abstract

Simultaneous localization and mapping (SLAM) is a very challenging yet fundamental problem in the field of robotics and photogrammetry, and it is also a prerequisite for intelligent perception of unmanned systems. In recent years, 3D LiDAR SLAM technology has made remarkable progress. However, to the best of our knowledge, almost all existing surveys focus on visual SLAM methods. To bridge the gap, this paper provides a comprehensive review that summarizes the scientific connotation, key difficulties, research status, and future trends of 3D LiDAR SLAM, aiming to give readers a better understanding of LiDAR SLAM technology, thereby inspiring future research. Specifically, it summarizes the contents and characteristics of the main steps of LiDAR SLAM, introduces the key difficulties it faces, and gives the relationship with existing reviews; it provides an overview of current research hotspots, including LiDAR-only methods and multi-sensor fusion methods, and gives milestone algorithms and open-source tools in each category; it summarizes common datasets, evaluation metrics and representative commercial SLAM solutions, and provides the evaluation results of mainstream methods on public datasets; it looks forward to the development trend of LiDAR SLAM, and considers the preliminary ideas of multi-modal SLAM, event SLAM, and quantum SLAM.

KEYWORDS
benchmark, LiDAR simultaneous localization and mapping (SLAM), loop closure detection, multi-sensor fusion, odometry

INTRODUCTION

Localization and navigation are crucial for acquiring space–time information, with significant implications for intelligent transport, security, economy, and people's well-being. The advent of 5G, the Internet of Things (IoT), and artificial intelligence has driven a soaring demand for precise space–time data in various societal and industrial contexts. Current localization and navigation technology relies on the Global Navigation Satellite System (GNSS), but its performance is hindered by weak satellite signals, rendering it highly unreliable in GNSS-denied settings (Shi et al., 2020), like indoors, jungles, and tunnels. In contrast, simultaneous localization and mapping (SLAM) technology is an active localization technology that operates independently of external signals, including satellites and geomagnetism. SLAM facilitates seamless navigation in both indoor and outdoor environments. Consequently, SLAM is essential, serving as the foundation and assurance for enabling intelligent perception in unmanned systems.

SLAM is a technology that simultaneously restores the pose of an unmanned system and constructs a map model of the surrounding environment (Cadena et al., 2016). When navigating in an unknown environment, an unmanned system acquires scene observations (e.g., images, point clouds, inertial data, etc.) through mounted sensors, and SLAM restores the real-time position and attitude of the unmanned system through spatial–temporal correlation of the observations, thereby realizing localization and mapping. SLAM methods are mainly divided into two categories according to the sensor type, that is, visual SLAM (Campos et al., 2021; Davison et al., 2007; Engel et al., 2014, 2018; Forster et al., 2014; Klein & Murray, 2007; Mur-Artal et al., 2015; Mur-Artal & Tardós, 2017) and LiDAR SLAM (Borrmann et al., 2008; Bosse & Zlot, 2009; Duan et al., 2023; Li et al., 2016; Li, Shi, et al., 2023a; Shan et al., 2020; Shan & Englot, 2018; Xu, Cai, et al., 2022; Zhang & Singh, 2014). Compared with visual SLAM, LiDAR SLAM (mainly referred to as 3D LiDAR) has the advantages of high mapping accuracy, good stability, unaffected by illumination changes, and no scale drift, which makes it more popular in real large-scale applications. In addition, the multi-sensor fusion scheme based on LiDAR SLAM has shown higher localization accuracy and stability, which is the current research hotspot and development trend. In short, the development of LiDAR SLAM technology will largely promote the progress and breakthroughs of intelligent perception applications such as 3D surveying and mapping, unmanned driving, military navigation, and deep space exploration.

In the past decades, SLAM research has made significant progress in both accuracy and robustness, and a number of milestone methods have emerged. However, due to factors such as motion distortions, occlusions, illumination changes, repeated/weak textures, and long-term mapping, SLAM technology remains challenging. To encourage beginners to engage in this field and inspire future research, scholars have summarized a series of excellent tutorials and reviews, which sort out the taxonomy, difficulties, benchmarking, and open problems of SLAM (Bailey & Durrant-Whyte, 2006; Cadena et al., 2016; Durrant-Whyte & Bailey, 2006; Fraundorfer & Scaramuzza, 2012; Grisetti et al., 2010; Scaramuzza & Fraundorfer, 2011; Yousif et al., 2015). However, they primarily focus on the visual SLAM problem, necessitating a comprehensive and systematic survey of LiDAR SLAM methods to bridge this gap. This paper summarizes the scientific connotation, key difficulties, research status, and future trends of LiDAR SLAM technology. Figure 1 shows a detailed methodological taxonomy of existing methods for LiDAR SLAM. The main contributions of this survey are four-fold:

- We present a systematic LiDAR SLAM review that covers the framework, challenges, taxonomy, benchmarking, future trends, etc.
- We give an in-depth overview of LiDAR SLAM methods, with brief summaries of advantages and limitations for each subcategory.
- We summarize commonly used datasets, evaluation metrics, and successful commercial SLAM solutions, and provide comprehensive comparisons of existing methods.
- We discuss the open problems and look forward to the new development trends to provide insightful guidance for the community.

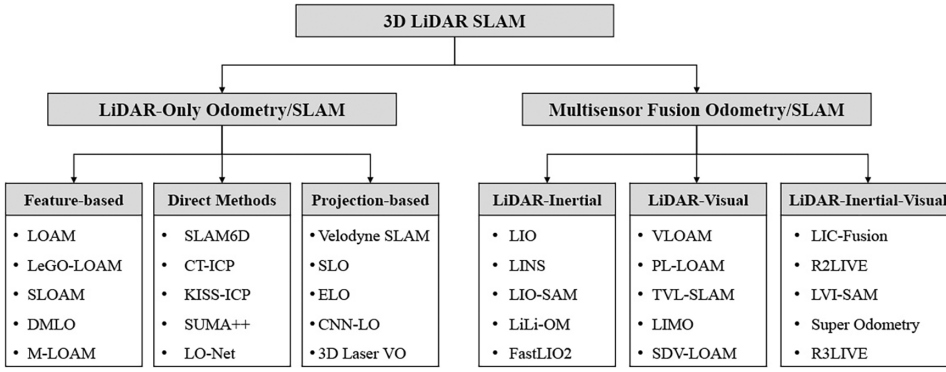


FIGURE 1 Methodological taxonomy of existing LiDAR SLAM. Mainstream methods are divided into LiDAR-only frameworks (feature-based, direct and projection-based) and multi-sensor fusion solutions (LiDAR-inertial, LiDAR-visual and LiDAR-inertial-visual).

The structure of this paper is as follows. Section 2 introduces the background of LiDAR SLAM, including its concept, characteristics, framework, and key difficulties. Section 3 presents an in-depth overview of current LiDAR-only and multi-sensor fusion methods. Section 4 summarizes publicly available datasets, evaluation metrics, and benchmarking results. Section 5 presents several promising future directions. Finally, Section 6 concludes the paper.

BACKGROUND

Anatomy of a LiDAR SLAM

A complete LiDAR SLAM system mainly includes two modules: the front end and the back end (Cadena et al., 2016). The front end (also known as odometry) first achieves data association and pose estimation of adjacent LiDAR scans through matching (a LiDAR scan refers to a collection of scanned points in a few moments of time, similar in concept to an image frame), and then uses loop-closure detection technology to construct pose constraints between the current LiDAR scan and the set of historical LiDAR scans. It can be seen that it contains two major components: scan matching and loop-closure detection. The back end uses optimization technology to globally adjust the model built by the front end to overcome the error cumulative introduced by incremental scan matching, which is an adjustment process. Figure 2 summarizes the classic framework and three components, that is, scan matching, loop-closure detection, and optimization are briefly reviewed.

Front end

Scan matching

It also known as point cloud registration is the problem of aligning two LiDAR scans in order to recover their relative poses. It is the de-facto standard technology for LiDAR data association, which is mainly divided into two categories: feature matching methods and direct matching methods.

Direct matching methods directly use LiDAR scan points for data association and pose estimation. Methods based on probability models such as normal distributions transform (NDT) (Biber & Strasser, 2003; Hong & Lee, 2017) have poor predictability of matching results, and correlation matching methods based on grid occupancy such as cartographer (Hess et al., 2016) have high computational complexity. Therefore, methods

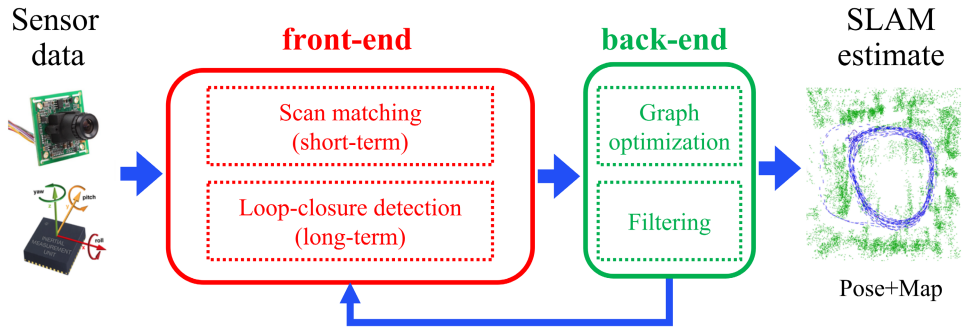


FIGURE 2 Framework of LiDAR SLAM (Cadena et al., 2016). A traditional system comprises two key components: a front end (red) and a back end (green). It takes sensor data as input and yields mapping and pose information.

based on geometric models such as iterative closest point (ICP) (Besl & McKay, 1992) have gradually become mainstream. Representative ICP-based SLAM methods include SLAM6D (Borrmann et al., 2008), IMLS-SLAM (Deschaud, 2018), GLIM (Koide et al., 2021a), CT-ICP (Dellenbach et al., 2022), Fast-LIO (Xu & Zhang, 2021), etc. ICP is a technique to simultaneously calculate the current optimal correspondences and the optimal rigid-body transformation, which has been widely used in point cloud registration. After more than 30 years of development, many mature variants of ICP have emerged. These variants improve one or several steps of the original ICP framework, including subset sampling (e.g., random sampling, octree sampling [Schnabel & Klein, 2006], Voxel-grid filtering [Rusu & Cousins, 2011], histogram sampling [Ervan & Temeltas, 2023]), distance metrics (e.g., point-to-line [Censi, 2008], point-to-plane [Chen & Medioni, 1992], symmetric point-to-plane [Rusinkiewicz, 2019], plane-to-plane [Koide et al., 2021b; Segal et al., 2009]), outlier rejection (e.g., sparse norm [Bouaziz et al., 2013; Li, Hu, & Ai, 2020a], anisotropy ICP [Maier-Hein et al., 2011], M-estimation [Chetverikov et al., 2005; Li, Hu, Ai, & Wang, 2021; Li, Hu, et al., 2022; Zhang et al., 2022]), and computational efficiency (e.g., Fast ICP [Zhang et al., 2022], EfficientVarICP [Rusinkiewicz & Levoy, 2001], and Anderson-accelerated ICP [Pavlov et al., 2018]). Although these methods have high accuracy, they are highly dependent on the initializations due to partial overlapping, occlusions, and noise, which make it easy to fall into local extrema. With the great success of deep learning, point cloud registration methods based on end-to-end neural networks have also attracted much attention (Aoki et al., 2019; Choy et al., 2020; Huang et al., 2021; Wang & Solomon, 2019a, 2019b). However, their performances in SLAM are not as good as the ones of traditional geometric model methods.

Feature matching methods perform data association and pose estimation based on highly significant features in LiDAR scans to improve computational efficiency. Among them, LOAM (Zhang & Singh, 2014) is the most classic and widely used method. It extracts both linear and planar features based on the curvature information of the LiDAR scan lines. To make LOAM suitable for various types of LiDAR, LOAM-Livox (Lin & Zhang, 2020) designed a feature selection strategy for petal-shaped scanning; MULLS (Pan et al., 2021) extracts ground, elevation, columnar, and linear features based on principal component analysis (PCA). These methods (Lin & Zhang, 2020; Pan et al., 2021; Shan & Englot, 2018; Wang, Wang, Chen, & Xie, 2021; Zhang & Singh, 2014; Zhao et al., 2022) establish correspondences by searching closest points, which is consistent with the ICP-type methods in essence, and also has the problem of local extrema. Existing schemes usually assume that the robot is in a local uniform or low-speed motion state, and use the pose prior information as the initializations to eliminate this problem.

Methods based on 3D feature descriptors can solve the above-mentioned prior hypothesis dependence problem. Similar to image feature matching (Li, Hu, & Ai, 2020b; Li, Shi, et al., 2023b; Li, Xu, et al., 2022; Lowe, 2004), 3D feature matching also includes the following steps: feature detection (e.g., intrinsic shape signatures [ISS] [Zhong, 2009], KeypointNet [You et al., 2020], USIP [Li & Lee, 2019], Rskdd-net [Lu et al., 2020], etc.), feature description (e.g., fast point feature histogram (FPFH), 3DMatch, SpinNet, GeDI, etc.), matching relationship

establishment (e.g., nearest neighbour distance ratio [Lowe, 2004], chi-square test [Zhong, 2009], etc.), and mismatch elimination (e.g., RANSAC [Barath et al., 2022; Fischler & Bolles, 1981; Li et al., 2017; Li, Hu, & Ai, 2020c], robust estimators [Jiang et al., 2023; Li, Hu, & Ai, 2021; Li, Zhao, et al., 2020; Yang et al., 2021], GORE [Li, 2022; Parra Bustos & Chin, 2018], etc.). However, due to problems such as disordered organization, uneven density, lack of texture, and structural occlusion, the outlier rate of 3D feature matching can reach 90%; in addition, descriptor-based methods have high computational complexity and are difficult to process in real time. These issues prevent them from becoming mainstream scan-matching methods for LiDAR SLAM.

Loop closure detection (LCD)

It establishes global constraints in the pose graph by identifying historical scenes highly similar to the current point cloud frame. Ensuring a robust LCD module is pivotal for overall system performance, as it effectively mitigates trajectory drift and facilitates high-precision map creation.

Current mainstream LCD techniques primarily encompass position prior-based, scene descriptor-based, and deep learning-based approaches. Position prior-based methods (Dellenbach et al., 2022; Shan & Englot, 2018) reduce interference frames via local windows, cutting down the search space and enhancing the practicality of point cloud matching-based LCD. However, they demand higher current position accuracy. Scene descriptor-based methods encode point cloud frames as feature vectors and detect closed loops by measuring feature vector similarity, enabling significant data compression and detection efficiency improvement. Notable techniques include M2DP (He, Wang, & Zhang, 2016), SC (Kim & Kim, 2018), LiDAR-IRIS (Wang et al., 2020), SC++ (Kim et al., 2022), etc. Deep learning-based methods excel in image/point cloud tasks and have been widely integrated into LCD, for example, LocNet (Yin et al., 2018), OREOS (Schaupp et al., 2019), and OverlapNet (Chen, Läbe, et al., 2022), yet exhibit limited generalization capabilities. However, LCD encounters two significant challenges. (1) Scene perception ambiguity: discriminating between scenes with high structural similarity (e.g., stairs, corridors, tunnels) remains challenging; and (2) insufficient information richness: Point cloud data lacks texture compared to images, making scene description more challenging. These issues can lead to incorrect closed loops, and even valid closed loops can be overshadowed in complex scenarios, posing significant challenges for subsequent graph optimization.

Recently, several works have explored innovative solutions for addressing the LCD problem. Shi et al. (2021; Ye et al., 2020) utilize indoor walls as a similarity criterion for scene comparison and employ wall matching for relative pose estimation, but this approach is limited in scenes lacking walls. Several works (Chen, Vizzo, et al., 2021; Lu, Zhou, et al., 2019; Schmiedel et al., 2015; Shi, Li, & Zhang, 2023a, 2023b; Wiesmann et al., 2022) aim for precise global vehicle localization within high-precision (HD) maps. However, offline maps typically entail higher memory consumption and necessitate timely updates. Some researchers incorporate transformers (Ma et al., 2022), semantics (Li, Kong, Zhao, Li, et al., 2021), and sequence data (Yin et al., 2022) into loop detection, but these methods demand extensive computational resources and warrant improvements in generalization capabilities. In summary, LCD remains an unsolved challenge in LiDAR SLAM, and achieving faster efficiency and higher accuracy recognition represents a complex and promising direction for future research.

Back end

The back end can be divided into filtering-based methods and graph optimization methods (Cadena et al., 2016). Filtering-based methods are early approaches for SLAM problems, which are derived from Bayesian estimation theory. It turns out that they are not as accurate and efficient as graph optimization methods in large-scale scenarios. Therefore, this article focuses on the currently more favoured graph optimization methods. More detailed information on filtering-based methods can be found in the surveys (Bailey & Durrant-Whyte, 2006; Durrant-Whyte & Bailey, 2006).

Among graph optimization methods, the SLAM problem is often formulated as maximum a posteriori (MAP) estimation, and factor graphs or pose graphs are often used to reason about the interdependencies between

variables. The pose graph is a graph structure composed of nodes and edges, where the nodes reflect the pose information of the robot at each moment, and the spatial constraints of poses at different moments are encoded in the edges (Grisetti et al., 2010). The goal of graph optimization is to find the optimal pose configuration that satisfies all constraints through adjustment. Iterative non-linear optimization, such as Gauss–Newton methods (Kaess et al., 2008; Kümmerle et al., 2011; Nasiri et al., 2021) and gradient descent methods (Grisetti et al., 2009; Olson et al., 2006), exploits the sparsity of the pose graph to enable the use of fast linear solvers, which has become the standard method of graph optimization technology and has developed many mature open-source tools, such as GTSAM (Dellaert, 2012), iSAM (Kaess et al., 2008, 2012), G2O (Kümmerle et al., 2011), Ceres (Agarwal et al., 2022), etc. However, iterative-based optimization methods cannot achieve the global optimal solution.

In recent years, researchers have found that under the condition of strong duality, the maximum likelihood estimation of this problem has a unique solution. Based on this observation, a series of convex relaxed optimization methods with globally optimal solutions are proposed (Carlone et al., 2016; Fan et al., 2020; Lajoie et al., 2019; Rosen et al., 2015). At present, these methods are still in the stage of theoretical research. In addition, scan matching and loop-closure detection inevitably result in outlier observations, while the above methods assume that the observation noise conforms to a normal distribution and is very sensitive to outliers. A large number of methods have introduced robust estimation technology to automatically detect outliers, such as l_1-norm estimation (Carlone et al., 2014; Casafranca et al., 2013), max-mixture model (Olson & Agarwal, 2013; Wang & Olson, 2014), M-estimation (Carlone & Calafiore, 2018; Lajoie et al., 2019), etc. However, these methods are limited by theoretical bottlenecks and only consider the graph optimization problem at low outlier ratios (Li, Zhang, & Hu, 2021). Therefore, algorithms with the global optimal solution and higher robustness are one of the development trends of graph optimization technology.

Challenges of LiDAR SLAM

While image processing technology has matured, visual sensors like cameras face challenges in dynamic urban settings, highways, and supermarkets due to sensitivity to lighting changes, dynamic environments, and adverse weather conditions. LiDAR has gained significant research interest because of its capacity to offer rich 3D information, wide field of view (FOV), and rapid update rates. Some methods have employed point cloud registration (Ji & Singh, 2017; Shan & Englot, 2018; Wang, Wang, Chen, & Xie, 2021), image representation (Cho et al., 2020; Wang, Saputra, et al., 2019), transformer (Liu et al., 2023), semantic information (Li, Kong, Zhao, Li, et al., 2021), branch and bound theory (Hess et al., 2016), and multi-source data fusion (Lin & Zhang, 2022a; Zuo et al., 2019) to enhance LiDAR SLAM. Despite their reported performance improvements, LiDAR SLAM still confronts the following challenges:

Weather condition

Laser signals exhibit less attenuation and longer propagation distances on sunny days while decaying significantly in rainy and foggy weather. LiDAR SLAM in adverse weather conditions often experiences degradation or failure.

Sensor diversity

There are three primary LiDAR types: mechanical, solid-state, and phased-array, each with distinct imaging principles, FOV, frequencies, and point cloud densities. Current LiDAR SLAM algorithms are tailored to specific LiDAR types and lack universal applicability across different LiDAR categories.

LCD

There is an immediate demand for a robust, efficient, and precise LCD algorithm. Accurately identifying revisited places within a vehicle's trajectory is crucial for mitigating cumulative pose errors and enhancing mapping accuracy. Conversely, erroneous detection poses significant risks to the system.

Multi-sensor fusion

Synchronizing and accurately calibrating information between LiDAR and other sensors (e.g., cameras, IMU, radar, wheel encoders) is challenging due to their significantly different frequencies.

Module cooperation

The standard LiDAR SLAM framework comprises front end and back end components. However, a comprehensive unmanned autonomous system necessitates integration with perception, control, and planning modules, which entails complex inter-module collaboration.

System efficiency

A single LiDAR frame generally contains tens or hundreds of thousands of points, making point cloud feature extraction and matching time-intensive. If HD maps are involved, processing frames and maps may further increase time requirements.

Relation to previous surveys

Recent computer vision advancements have resulted in numerous reviews on visual SLAM (Bailey & Durrant-Whyte, 2006; Durrant-Whyte & Bailey, 2006; Fraundorfer & Scaramuzza, 2012; Nister et al., 2004; Saputra et al., 2018; Taketomi et al., 2017; Yousif et al., 2015). While these reviews have significantly advanced robotics and autonomous driving research, they have regrettably overlooked LiDAR technology. Cadena et al. (2016) extensively review the current state of SLAM and delve into potential future directions. Grisetti et al. (2010) present an introductory description of the graph-based SLAM and discuss state-of-the-art (SOTA) solutions based on least-squares error minimization. Debeunne & Vivet (2020) offer a comprehensive survey on visual-LiDAR SLAM. However, their sections discussing LiDAR are relatively limited in content. Shi, Zhang, and Li (2023) provide a comprehensive review of place recognition methods employing LiDAR sensors and conduct an in-depth review of related research, while it merely represents a module of LiDAR SLAM. Huang et al. (2021) summarize common LiDAR SLAM frameworks and core module functions but provide an insufficient literature review.

Unlike prior research, our review exclusively concentrates on LiDAR-based SLAM studies. It delivers a comprehensive overview of the framework, task challenges, method taxonomy, and code resources. Furthermore, it collects more open-source datasets, evaluation metrics, commercial SLAM solutions and experimental performance on public datasets while pointing out promising future research directions.

LITERATURE REVIEW

LiDAR-only SLAM

Feature-based methods

LiDAR actively emits and receives laser signals to create 3D point clouds depicting real-world surfaces. However, dense point clouds often carry redundant data, impeding computation efficiency. Feature-based systems aim to extract representative features with tangible physical significance and estimate vehicle motion through feature association. Typically, feature-based methods encompass basic and advanced features, depending on their characteristics. Table 1 gives a summary of these methods. Figures 3 and 4 illustrate the workflow and mapping results of various representative methods.

Basic feature-based SLAM

Basic features refer to fundamental geometric properties like coordinates, colours, intensities, curvatures, and normals. They can be derived from individual or grouped points to comprehend the underlying 3D scene. One type of feature pertains to appearance characteristics, encompassing colour, reflectance, and texture, which describe the visual appearance of the point cloud and aid in material differentiation on surfaces. As shown in Figures 3a and 4a, LOAM-Livox (Lin & Zhang, 2020) introduces a SLAM system tailored to LiDARs with limited FOV, building upon the basic structure of standard LOAM while bringing forth notable innovations, including reflectivity-driven point selection, iterative pose optimization, and parallelization. Intensity-SLAM (Wang, Wang, & Xie, 2021) devises an intensity-assisted framework utilizing both intensity and geometry data for odometry estimation, integrating intensity-based LCD and factor graph optimization, and outperforming geometry-only methods across varied environments.

The other common types, namely geometric features, encompassing coordinates, distances, normals, curvature, and density, which describe point properties related to their spatial arrangement or geometry. LOAM (Ji & Singh, 2017; Zhang & Singh, 2014) is a pioneering LiDAR SLAM framework that extracted 3D planar and edge points using curvatures, enabling high-frequency odometry calculations through point-to-line and point-to-plane distance minimization and linear interpolation-based motion compensation. It optimizes poses by registering accumulated scans into maps, achieving low-frequency mapping and transformation integration. Subsequently, several LOAM variants incorporate LCD (Rozenberszki & Majdik, 2020; Shan & Englot, 2018), local bundle adjustment (Liu & Zhang, 2021), and non-linear distortion compensation (BA) (Wang, Wang, Chen, & Xie, 2021) to bolster system robustness. Other variants (Jiao et al., 2021, 2022; Lin et al., 2020) have successfully extended to multi-robot systems. Another research branch employs learning techniques to detect interest or matching points. CAE-LO (Yin et al., 2020) employs a convolutional auto-encoder-based odometry for detecting interest points from spherical ring data and a 3D encoder for feature extraction from multi-resolution voxel models. It ensures efficient computation, retains the original 3D shape, and exhibits high versatility. DMLO (Li & Wang, 2020) devises a sparse matching odometry framework involving point cloud projection onto a cylindrical plane, cascade convolutional neural network (CNN) for grid-wise feature extraction and corresponding points, and rigid transformation solution through singular value decomposition (SVD).

Advanced feature-based SLAM

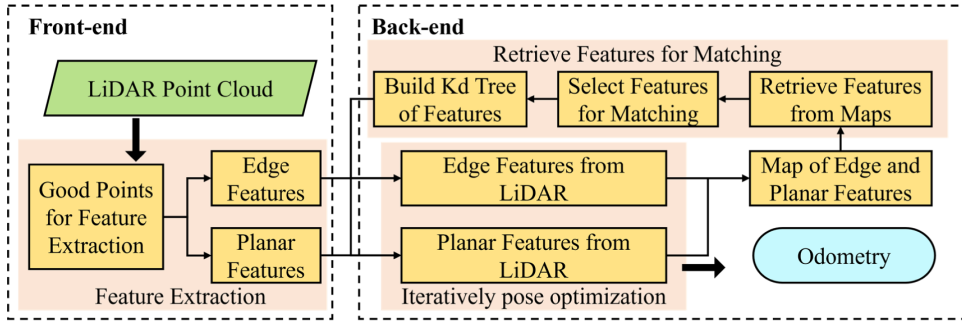
Advanced features encompass broader attributes derived through sophisticated processing techniques to describe the scene and object-level information. The first type is semantics, involving the meaningful interpretation of points based on their content or 3D context. By assigning labels (e.g., car, road, building, and tree) to represent point categories, the understanding of elements within the 3D environment will be enhanced. SA-LOAM (Li, Kong, Zhao, Li, et al., 2021) (Figure 3b) introduces a semantic-enhanced SLAM, building upon F-LOAM (Wang,

TABLE 1 Summary of feature-based SLAM. These methods are organized based on feature types, method names, published years and code websites.

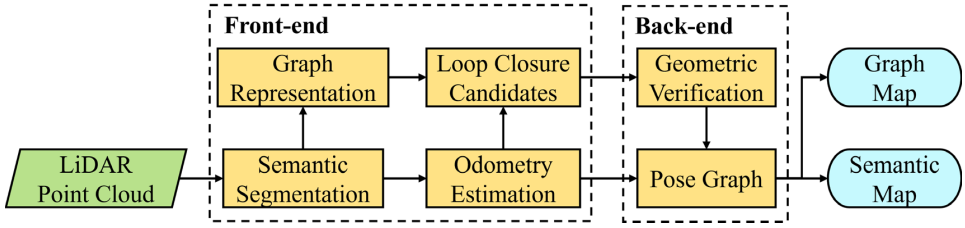
Type	Feature	Method	Year	Code
Basic feature-based SLAM	Appearance	LOAM-Livox (Lin & Zhang, 2020)	2020	https://github.com/hku-mars/loam_livox
		Intensity-SLAM (Wang, Wang, & Xie, 2021)	2021	https://github.com/wh200720041/intensity_slam
	Geometry	LOAM (Zhang & Singh, 2014; Ji & Singh, 2017)	2017	https://github.com/laboshinl/loam_velodyne
		LeGO-LOAM (Shan & Englot, 2018)	2018	https://github.com/RobustFieldAutonomyLab/LeGO-LOAM
		LOL (Rozenberszki & Majdik, 2020)	2020	https://github.com/RozDavid/LOL
		Lin et al. (Lin et al., 2020)	2020	https://github.com/hku-mars/decentralized_loam
		CAE-LO (Yin et al., 2020)	2020	https://github.com/SRainGit/CAE-LO
		DMLO (Li & Wang, 2020)	2020	×
		BALM (Liu & Zhang, 2021)	2021	https://github.com/hku-mars/BALM
		F-LOAM (Wang, Wang, Chen, & Xie, 2021)	2021	https://github.com/wh200720041/floam
		M-LOAM (Jiao et al., 2021)	2021	https://ram-lab.com/file/site/m-loam
Advanced feature-based SLAM	Semantics	S-LOAM (Chen et al., 2020)	2020	×
		PSF-LO (Chen, Wang, et al., 2021)	2021	×
		SA-LOAM (Li, Kong, Zhao, Li, et al., 2021)	2021	×
		Generalized-LOAM (Honda et al., 2022)	2023	https://github.com/kohonda/proj-gloam
	Shape	Bosse et al. (2009)	2009	×
		Velas et al. (2016)	2016	https://github.com/robofit/but_velodyne_lib
		Grant et al. (2019)	2019	×
		MULLS (Pan et al., 2021)	2021	https://github.com/YuePanEdward/MULLS

Wang, Chen, & Xie, 2021) (Figure 4b), with a semantic-assisted ICP and a semantic graph-based LCD. Leveraging semantics elevates accuracy in localization, enhances LCD, and ensures global map consistency, even in expansive environments. Other works leverage semantic segmentation techniques to offer valuable insights for local shape description (Honda et al., 2022), odometry estimation (Chen, Wang, et al., 2021) (Figure 4c), and optimization (Chen et al., 2020).

Shape characteristic is another vital advanced feature that describes and elucidates the overall shape and structure of surfaces, facilitating the analysis of shape-related aspects within the 3D environment. Bosse and

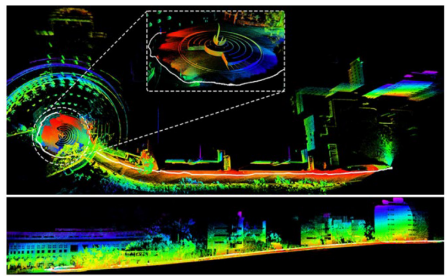


(a) Basic Feature-based SLAM

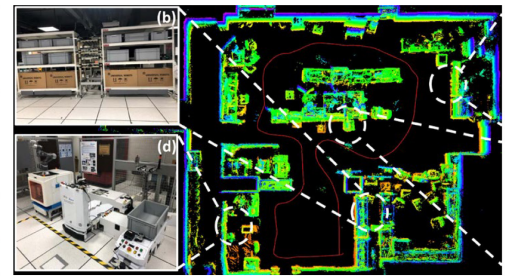


(b) Advanced Feature-based SLAM

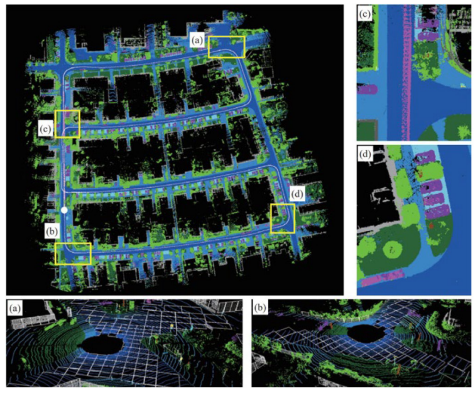
FIGURE 3 Four representative feature-based methods. (a) and (b) are originally shown in Lin & Zhang (2020) and Li, Kong, Zhao, Li, et al. (2021), respectively.



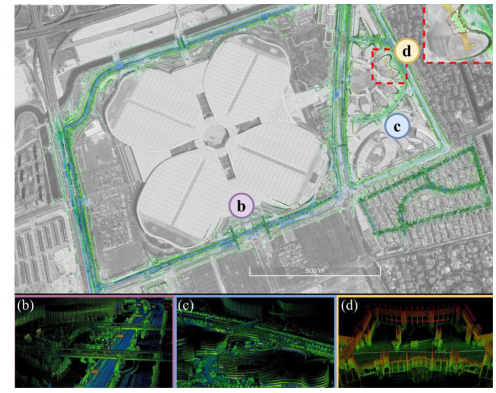
(a) LOAM-Livox [56] in Campus



(b) F-LOAM [58] in Warehouse



(c) PSF-LO [134] in Urban



(d) MULLS [57] in Urban

FIGURE 4 Mapping results of representative feature-based methods.

Zlot (2009) incorporate 3D grid-based local geometry and vehicle motion constraints into point matching, laying the groundwork for a thorough 3D SLAM and eliminating the necessity for odometric or inertial sensors. Velas et al. (2016) form a line cloud representation via randomly generated collar line segments (CLS) and estimate a transformation aligning matched line segments onto a 3D plane, addressing LiDAR point cloud sparsity. Grant et al. (2019) employ an efficient plane detector for stable feature extraction, serving both localization and graph-based SLAM landmarks. This method accommodates non-smooth trajectories, making it applicable to humanoid and aerial robots without necessitating odometry measurements. As shown in Figure 4d, MULLS (Pan et al., 2021) utilizes dual-threshold ground filtering and PCA for extracting ground, facade, pillar, and beam features. It estimates ego-motion via optimized point-to-point (plane, line) error metrics within each point class and employs hierarchical pose graph optimization to mitigate drift from dead reckoning.

Direct methods

Different from feature-based methods, direct LiDAR SLAM operates on the entire point cloud, eliminating the need for feature extraction and utilizing direct matching for pose calculation. This characteristic renders it well-suited for featureless environments and provides higher accuracy and resolution maps. Contingent on the matching process, direct SLAM methods typically encompass two primary matching processes: scan-to-scan and scan-to-model methods. Table 2 gives a method classification while Figures 5 and 6 shows several representative direct methods.

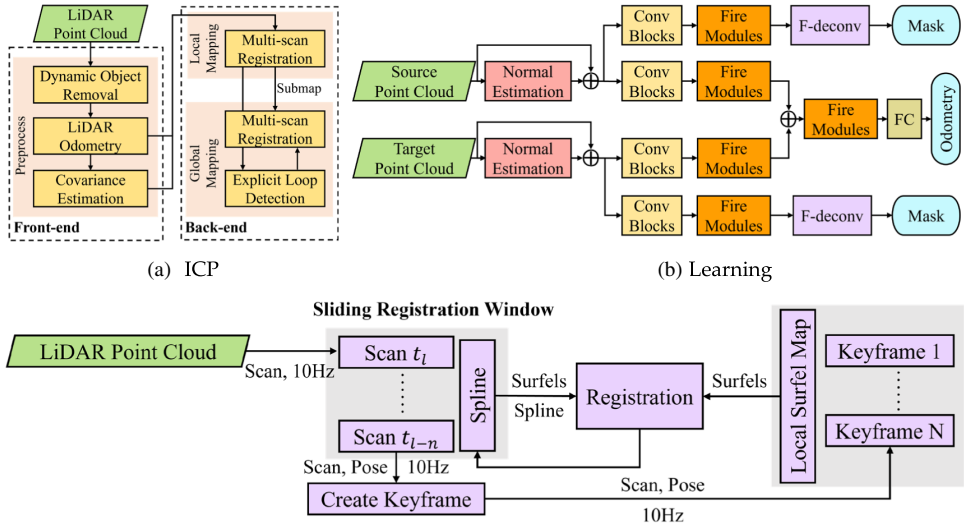
Scan-to-scan matching

Scan-to-scan methods typically employ point cloud matching to compute the relative pose between two sequential point clouds, incrementally building a point cloud map. A prevalent approach involves employing ICP (Besl & McKay, 1992) for ascertaining the relative pose between two point clouds. ICP, a well-established point cloud registration, iteratively refines transformation parameters by identifying the nearest corresponding points. It finds broad application in robotics, 3D reconstruction, and geographic information systems (GIS). SLAM6D (Borrmann et al., 2008) pioneers the application of ICP for sequential scan registration, using distance criteria within a graph to detect closed loops. This method effectively mitigates error accumulation issues in sequential matching approaches. The original ICP excels in achieving precise point cloud alignment and generating dense maps, rendering it suitable for accurate localization and mapping. However, it heavily depends on the initial pose and struggles in high-speed or occluded scenarios. Subsequent research introduces ICP variants like generalized ICP (GICP) (Koide et al., 2021a; Reinke et al., 2022), KL-divergence-based ICP (Yokozuka et al., 2021) (Figure 6a), and voxelized generalized ICP (VGICP) (Frosi & Matteucci, 2022) for the system.

Inspired by deep learning's success in visual odometry, researchers have extended this technology to LiDAR odometry. The network takes two real-time point clouds as input and predicts the odometry pose. As shown in Figure 5(b), LO-Net (Li et al., 2019) introduces a scan-to-scan LiDAR odometry network, improving feature representation with a novel mask-weighted geometric constraint loss. It implicitly handles sequential dependencies and data dynamics. Furthermore, it includes a scan-to-map module that enhances accuracy by incorporating geometric and semantic information. DeepVCP (Lu, Wan, et al., 2019) introduces an end-to-end framework for point cloud registration. It generates corresponding points using learned matching probabilities among candidate groups, which enhances registration accuracy and effectively mitigates interference from dynamic objects. HPPLO-Net (Zhou et al., 2023) (Figure 6b), an unsupervised hierarchical framework for large-scale dynamic outdoor scenes, achieves precise pose estimation using a differentiable point-to-plane solver with scene flow information. It employs a differentiable weighted point-to-plane SVD to resolve the pose matrix and rectify data association inaccuracies. Learning-based odometry excels at discerning intricate patterns in LiDAR data, resulting in enhanced accuracy. Nevertheless, it lacks transparency and interpretability, posing difficulties in comprehending the model's decision-making process and hindering generalization, particularly in scenarios significantly divergent from the training data distribution.

TABLE 2 Summary of direct LiDAR SLAM. These methods are organized based on matching types, method names, published years and code websites.

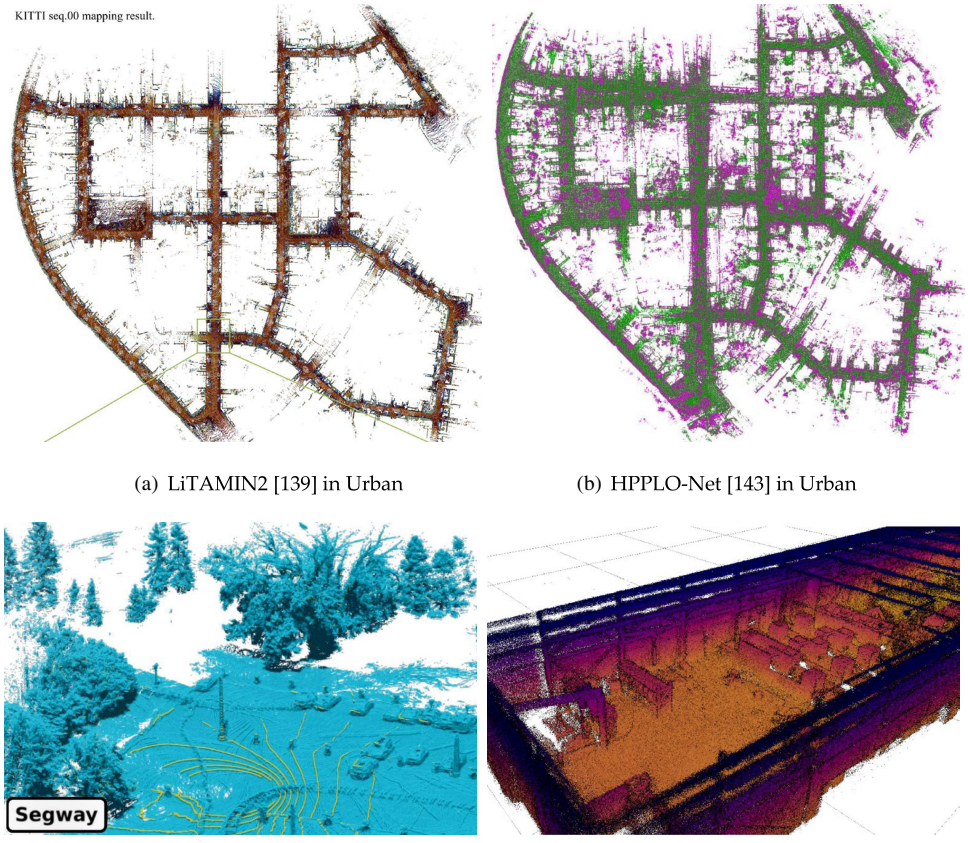
Type	Feature	Method	Year	Code
Scan-to-scan	Handcrafted ICP	SLAM6D (Borrmann et al., 2008)	2008	×
		GLIM (Koide et al., 2021a)	2021	×
		LiTAMIN2 (Yokozuka et al., 2021)	2021	https://github.com/bzdfzfer/litamin2
		LOCUS 2.0 (Reinke et al., 2022)	2022	https://github.com/NeBula-Autonomy/LOCUS
	Neural network	LO-Net (Li et al., 2019)	2019	https://github.com/PawitKoch/LO-Net-pytorch/tree/main
		DeepVCP (Lu, Wan, et al., 2019)	2019	https://github.com/vccheng2001/DeepVCP-Pointcloud-Registration
		HPPLO-Net (Zhou et al., 2023)	2023	https://github.com/IMRL/HPPLO-Net
Scan-to-map	Point cloud map	Cartographer (Hess et al., 2016)	2016	https://github.com/googlecartographer
		HDL-Graph-SLAM (Koide et al., 2019)	2019	https://github.com/koide3/hdl_graph_slam
		CT-ICP (Dellenbach et al., 2022)	2022	https://github.com/jedeschaud/ct_icp
		KISS-ICP (Vizzo et al., 2023)	2023	https://github.com/PRBonn/kiss-icp
	Surfel map	Elastic LiDAR (Park et al., 2018) fusion	2018	×
		Droeschel & Behnke (2018)	2018	×
		SuMa (Behley & Stachniss, 2018)	2018	https://jbehley.github.io/projects/surfel_mapping
		SuMa++ (Chen et al., 2019)	2019	https://github.com/PRBonn/semantic_suma
		MARS (Quenzel & Behnke, 2021)	2021	https://github.com/AIS-Bonn/lidar_mars_registration
	Other types	IMLS-SLAM (Deschaud, 2018)	2018	×
		Vizzo et al. (2021)	2021	https://github.com/PRBonn/puma



(a) ICP (b) Learning

(c) Surfel Map

FIGURE 5 Three representative direct methods. (a–c) are originally shown in Koide et al. (2021a), Li et al. (2019) and Quenzel & Behnke (2021), respectively.



(a) LiTAMIN2 [139] in Urban (b) HPPLO-Net [143] in Urban

(c) KISS-ICP [145] in Campus (d) MARS [150] in Indoor

FIGURE 6 Mapping results of representative direct methods.

Scan-to-model matching

Scan-to-model methods align acquiring scans with online submaps or pre-existing maps. While models offer richer surface information than sparse individual scans, significant quantitative disparities pose challenges for the matching process. A prevalent model is the point cloud map, constructed directly from accumulated 3D points, storing spatial coordinates, colour, intensity, or other attributes, ideal for precise modelling. Cartographer (Hess et al., 2016) employs a branch-and-bound approach (Nicolai et al., 2016a) for scan-to-submap matching as loop closure constraints, enabling real-time mapping of vast areas, even up to tens of thousands of square metres, with optimized results for the operator. HDL-Graph-SLAM (Koide et al., 2019) employs Graph SLAM (Grisetti et al., 2010) to construct an offline map and combines NDT (Biber & Strasser, 2003) with angular velocity-based pose prediction using an unscented Kalman filter (UKF) (Wan & Van der Merwe, 2000) for tracking, facilitating long-term and wide-area behaviour assessment. CT-ICP (Dellenbach et al., 2022) integrates scan continuity and discontinuity between scans by solving a scan-to-map registration, enabling elastic scan distortion during registration for higher precision and enhanced robustness against high-frequency motion. As depicted in (Figure 6c), KISS-ICP (Vizzo et al., 2023) combines point-to-point ICP with adaptive thresholding, a robust kernel, universal motion compensation, and point cloud subsampling, creating a low-parameter system compatible with various sensor types.

The surface element (surfel) map is another well-received model, with each surfel encoding information about a segment of the environment's surface, including position, radius, orientation, timestamps, and attributes such as colour or reflectivity. Surfels provide a more space-efficient alternative to storing raw 3D points, albeit with increased computational demands. They also bolster map accuracy by reducing noise and uncertainty, thus enhancing localization precision. Elastic LiDAR fusion (Park et al., 2018) employs a linear continuous-time trajectory to eliminate the need for global trajectory optimization. It introduces map deformation and improves the precision of the reconstructed dense map via probabilistic surfel fusion. Droeschel & Behnke (2018) aggregate local multi-resolution maps through surfel-based registration, forming a hierarchical graph structure treating individual 3D scans as subgraphs. It employs graph optimization to tackle drift and misalignment while enabling measurement interpolation between scan poses. SuMa (Behley & Stachniss, 2018), depicted in Figure 5c, utilizes frame-to-model ICP for aligning points between vertex maps and constructs a surfel-based map. It harnesses map representation to create a virtual map view, enhancing detection robustness even in scenarios with limited overlap. SuMa++ (Chen et al., 2019) further integrates semantic information (Milioto et al., 2019) to facilitate the mapping process. Multi-adaptive-resolution-surfel (MARS) (Quenzel & Behnke, 2021) depicted in Figure 6(d), integrates a continuous-time B-spline trajectory representation with a Gaussian mixture model (GMM) to align multi-resolution surfel maps collaboratively. Utilizing sparse voxel grids and permutohedral lattices ensures rapid access to map surfels, while an adaptive resolution selection scheme significantly enhances registration speed.

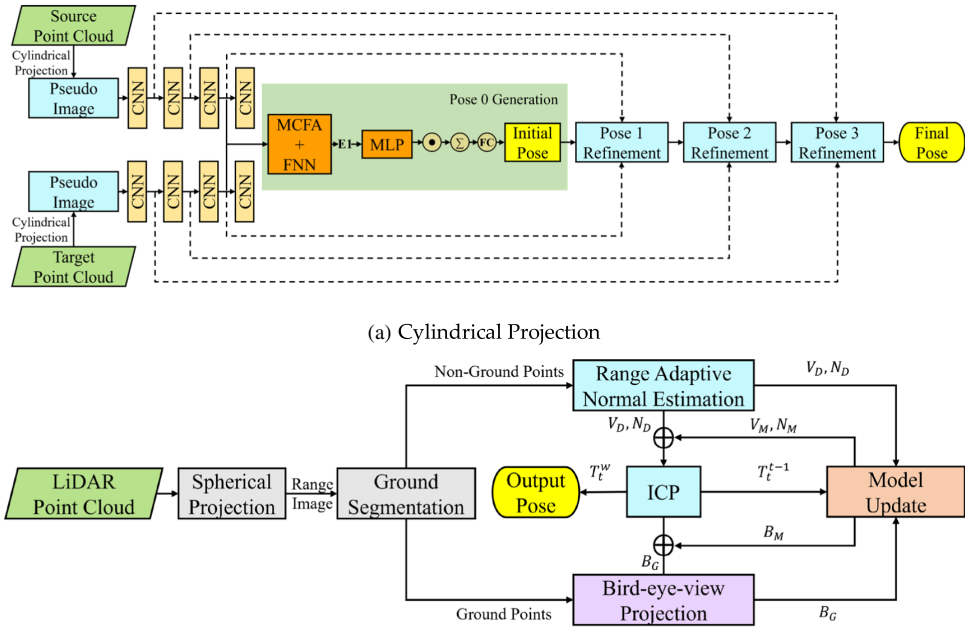
Other works strive to create innovative maps for enhanced localization and mapping accuracy. IMLS-SLAM (Deschaud, 2018) selects an implicit moving least square surface as a map representation. Vizzo et al. (2021) employ frame-to-mesh ICP, representing the map as a Poisson surface-reconstructed triangle mesh within a sliding window. It provides a 3D map with finer geometric details compared to traditional methods employing truncated signed distance functions (TSDF) (Millane et al., 2018) or surfels (Behley & Stachniss, 2018; Park et al., 2018).

Projection-based methods

Projection-based methods typically transform point clouds into images and employ visual odometry techniques to estimate the sensor's self-motion. A key advantage lies in the 2D representations (e.g., images or grids), which streamline data processing, reduce computational demands, and tap into well-established image processing techniques. Nonetheless, the projection process invariably leads to information loss, potentially compromising mapping precision by distorting or omitting fine environmental details. As summarized in Table 3, projection-based

TABLE 3 Summary of project-based LiDAR SLAM. These methods are organized based on projection types, method names, published years and code websites.

Projection type	Method	Year	Code
Cylindrical projection	3D Laser VO (Tong & Barfoot, 2013)	2013	×
	CNN-LO (Nicolai et al., 2016b)	2016	×
	DeepPCO (Wang, Saputra, et al., 2019)	2019	https://github.com/soup1997/DeepPCO
	SLO (Guadagnino et al., 2022)	2022	×
	TransLO (Liu et al., 2023)	2023	https://github.com/IRMVLab/TransLO
Spherical projection	Velodyne SLAM (Moosmann & Stiller, 2011)	2011	×
	Cho et al. (2020)	2020	×
	ELO (Zheng & Zhu, 2021)	2021	×
	ECTLO (Zheng & Zhu, 2022)	2022	×



(a) Cylindrical Projection

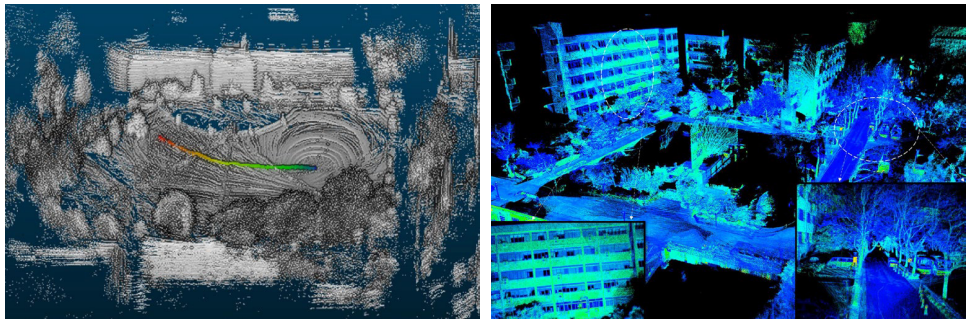
(b) Spherical Projection

FIGURE 7 Two representative projection-based methods. (a) and (b) are originally shown in Liu et al. (2023) and Zheng & Zhu (2021), respectively.

systems are classified as spherical and cylindrical based on the chosen projection method. Figures 7 and 8 illustrates corresponding two projection methods.

Cylindrical projection

Given a 3D point $p = (x, y, z)$ and corresponding image coordinate (u, v), the cylindrical projection Π_C is defined as equation (1). 3D Laser VO (Tong & Barfoot, 2013) introduces a 3D LiDAR-based visual odometry that generates



(a) CNN-LO [158] in Indoor (b) ECTLO [162] in Campus

FIGURE 8 Mapping results of representative feature-based methods.

intensity images, tracks sparse visual features, and mitigates motion distortion using the Gaussian process Gauss–Newton (GPGN) (Tong et al., 2012) for continuous-time state estimation. CNN-LO (Nicolai et al., 2016b), depicted in Figure 8a, leverage recent advancements in image classification to reduce the LiDAR scan's state space and employ standard fully connected layers to learn odometry estimation from the extracted features. DeepPCO (Wang, Saputra, et al., 2019) presents an end-to-end deep parallel neural network for accurate odometry by integrating 2D panoramic depth projections and two sub-networks. It eliminates traditional intermediate modules like scan matching and geometric estimation while performing well with various neural network architectures. SLO (Guadagnino et al., 2022) utilizes SuperPoint (DeTone et al., 2018) on intensity images for feature selection and introduces a self-supervised online fine-tuning procedure for the feature extraction network, which doesn't rely on labels or ground truth poses. As illustrated in Figure 7a, TransLO (Liu et al., 2023) designs the first transformer-based LiDAR odometry, which can process tens of thousands of points simultaneously by projecting points onto a 2D surface and then feeding them into a local transformer with linear complexity.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \Pi_C(x,y,z) = \begin{bmatrix} \lfloor atan2(y,x)/\Delta\theta \rfloor \\ \left\lfloor arcsin\left(z, \sqrt{x^2+y^2+z^2}\right)/\Delta\phi \right\rfloor \end{bmatrix} \tag{1}$$

where $\Delta\theta$ and $\Delta\phi$ is the average horizontal and vertical angle resolution between consecutive beam emitters, respectively. $\lfloor \cdot \rfloor$ means the floor computation. Generally, the element at (u, c) is filled with two-channel data (d, z), where $d = \sqrt{x^2 + y^2}$. The elements in 2D positions without projected 3D points are $(d, z) = (0, 0)$.

Spherical projection

Equation (2) describes the mapping function Π_S from the Cartesian to its corresponding spherical coordinate. Spherical projection maps point clouds onto a distortion-free 2D sphere, ideal for applications demanding a full 360° view. Velodyne SLAM (Moosmann & Stiller, 2011) is a specialized system tailored to the Velodyne laser scanner, focusing on flat surface measurements and incorporating an offline filtering step to enhance the generated map for detailed city mapping. Cho et al. (2020) introduce the first unsupervised learning-based odometry, utilizing input vertices and a geometry-aware consistency loss calculation to eliminate the need for time-consuming labelling procedures. ELO (Zheng & Zhu, 2021), depicted in Figure 7(b), exploits non-ground spherical range images and bird's-eye-view (BEV) maps for ground points, formulating odometry as a non-linear least squares minimization problem. It incorporates a range-adaptive technique for robustly estimating local surface normals, merging points and normals using a rapid memory-efficient model update scheme. As shown in Figure 8(b), ECTLO (Zheng & Zhu, 2022) develops an efficient odometry method for solid-state LiDARs, utilizing a point-to-plane GMM for registration, implementing a continuous-time motion model to mitigate distortions, and keeping all map points within a single range image to enable implicit data association in parallel.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \Pi_S(x, y, z) = \begin{bmatrix} \frac{1}{2}\left(1 - \frac{\theta}{\pi}\right) w_S \\ \left[1 - (\phi + f_{up}) f^{-1}\right] h_S \end{bmatrix} \tag{2}$$

where the sensor's vertical FOV $f = f_{up} + f_{down}$ comprises the upper (f_{up}) and lower (f_{down}) parts. w_S and h_S are the width and height of image, respectively. The azimuth angle $\theta = arctan(y / x)$, $-\pi < \theta \leq \pi$, while the elevation angle $\phi = arcsin\left(z / \sqrt{x^2 + y^2 + z^2}\right)$, $-\frac{\pi}{2} < \phi < \frac{\pi}{2}$.

Summary

In this section, LiDRA-only SLAM is categorized into three primary categories: feature-based, direct, and projection-based methods, further subdivide them into more subclasses based on common characteristics. Several key observations are summarized as follows:

- LiDAR sensors, unlike visual counterparts, are impervious to lighting variations and texture dependencies and offer precise depth data. LiDAR-only odometry remains the foremost choice for its reliability. The advent of groundbreaking odometry frameworks like LOAM (Ji & Singh, 2017) and LeGO-LOAM (Shan & Englot, 2018) has empowered single-sensor localization and mapping, catalysing the adoption of robotics and autonomous driving across civilian, commercial, educational, and military domains. This development has energised the robotics community, propelling innovation in resilient and cost-effective robotic solutions.
- Handcrafted SLAM frameworks (Hess et al., 2016; Shan & Englot, 2018), (Dellenbach et al., 2022; Ji & Singh, 2017; Lin & Zhang, 2020; Wang, Wang, & Xie, 2021) consist of well-established modules, serve as current mainstream solutions, and form a strong foundation for multi-sensor fusion. However, the system's maintenance and updates pose challenges, particularly in light of emerging sensors and hardware (Zheng & Zhu, 2022), necessitating ongoing algorithmic improvements. The rapid evolution of 3D computer vision has partly propelled learning-based odometry frameworks (Chen, Wang, et al., 2021; Li et al., 2019; Zhou et al., 2023), which primarily focus on feature extraction and matching rather than providing a complete solution. Despite yielding impressive results, enhancing their generalization capabilities (Chen et al., 2020; Li, Kong, Zhao, Li, et al., 2021) and managing their substantial computational demands (Liu et al., 2023) remains a priority.
- Selecting a suitable LiDAR SLAM method should align with specific requirements related to environments, sensors, mapping, memory, and platforms. Feature-based methods excel in structured environments, as they prioritize tracking distinctive features, making them advantageous and memory-efficient in urban settings with landmarks, pedestrians, and vehicles. However, their reliance on features renders them ineffective in feature-sparse, featureless, and dynamic environments. Direct methods preserve complete 3D information from point cloud data, ideal for precise mapping but computationally intensive. Conversely, projection-based techniques convert point clouds into images, enhancing efficiency, especially on resource-limited platforms. Nevertheless, projection entails information loss, and sensor tilts may distort projections, impacting map quality.

Multi-sensor fusion SLAM

LiDAR-inertial (LI) fusion

LiDAR usually has a low scanning frequency, and the perceived point clouds suffer from motion distortions. In the case of high-speed or complex motion, the accuracy of the linear distortion correction model is poor, which easily leads to the deterioration of the LiDAR-only SLAM methods. However, inertial sensors have high output

frequency and local accuracy, which can well compensate for the internal motion of the points inside a single LiDAR scan. LI fusion can effectively improve the robustness and accuracy of SLAM systems in featureless scenarios such as long corridors. Depending on how they are combined, LI SLAM systems can be classified into loosely coupled methods and tightly coupled ones. Table 4 provides a method summary of these methods.

TABLE 4 Summary of LiDAR-inertial fusion LiDAR SLAM. These methods are organized based on coupled types, method names, published years and code websites.

Type	Feature	Method	Year	Code
Loosely coupled	IMU odometry aided	IMU-aided LOAM (Zhang & Singh, 2014)	2014	https://github.com/HKUST-Aerial-Robotics/A-LOAM
		Kim et al. (Kim et al., 2023)	2023	×
	IMU odometry and LO fusion	PPP-LOAM (Li, Pei, et al., 2020)	2020	×
		LION (Tagliabue et al., 2021)	2021	×
		EKF-LOAM (Júnior et al., 2022)	2022	×
Tightly coupled	Filter-based	LINS (Qin et al., 2020)	2020	https://github.com/ChaoqinRobotics
		Fast-LIO (Xu & Zhang, 2021)	2021	https://github.com/hku-mars/FAST_LIO
		Fast-LIO2 (Xu, Cai, et al., 2022)	2022	https://github.com/hku-mars/FAST_LIO
		Faster-LIO (Bai et al., 2022)	2022	http://github.com/gaoxiang12/faster-lio
		Puma-LIO (Jiang & Shen, 2022)	2022	https://www.github.com/lewisjiang/puma-lio
		AdaLIO (Lim et al., 2023)	2023	×
		Swarm-LIO (Zhu et al., 2023)	2023	×
		HD-LIO (Wang et al., 2023)	2023	×
		Duan et al. (Duan et al., 2023)	2023	×
	Graph optimization	LIPS (Geneva et al., 2018)	2018	https://github.com/rpng/lips
		LIOM (Ye et al., 2019)	2019	https://ram-lab.com/file/hyye/lio-mapping
		LIO-SAM (Shan et al., 2020)	2020	https://github.com/TixiaoShan/LIO-SAM
		GR-LOAM (Su et al., 2021)	2021	×
		LiLi-OM (Li, Li, & Hanebeck, 2021)	2021	https://github.com/KIT-ISAS/lili-om
		LIO-vehicle (Chou & Chou, 2022)	2022	×
		D-LIOM (Wang et al., 2022)	2022	https://cslinzhang.github.io/D-LIOM/D-LIOM.html
		Koide et al. (2022)	2022	×
		DLIO (Chen et al., 2023)	2023	https://github.com/vectr-ucla/direct_lidar_inertial_odometry

Loosely coupled LI SLAM

Loose coupling typically processes LiDAR and IMU observations separately and then fuses their results. The simplest way is to use the IMU measurements to correct the non-linear motion distortions of LiDAR scans and obtain the IMU odometry through the IMU pre-integration, which is provided as initializations to the scan matching to improve its robustness. This coupling can be directly applied to LiDAR-only SLAM systems with little modification. Representative works are IMU-aided LOAM, LeGO-LOAM, MULLS, Adaptive Keyframe LIO (Kim et al., 2023) (Figure 10a), etc. Another more popular loose coupling approach is to perform state estimation based on LiDAR and IMU observations separately, and then fuse the LiDAR odometry (LO) and IMU odometry results using Kalman filtering (Kalman, 1960) or other technologies. For example, methods such as EKF-LOAM (Júnior et al., 2022) (Figures 9 and 10b) and PPP-SLAM (Li, Pei, et al., 2020) use extended Kalman filters (EKF) to integrate measurements from LiDAR, IMU, and GNSS during the state optimization stage. LION (Tagliabue et al., 2021) expresses the results of LO and IMU odometry as a factor graph and optimizes it through a fixed-lag sliding window smoother. In loose coupling, the separation of scan matching and result fusion reduces the computational load, making it highly efficient, easy to implement, and easy to expand. However, because loose coupling ignores other state quantities of the system (such as velocity), it may lead to information loss, so that the advantages of each sensor cannot be maximized.

Tightly coupled LI SLAM

Tight coupling formulates the state estimation factors of LiDAR and IMU observations as a joint optimization problem. Since tightly coupled methods overcome the above-mentioned limitations, they show better performance in complex scenarios and have become the mainstream methods for LI fusion. Tightly coupled methods can

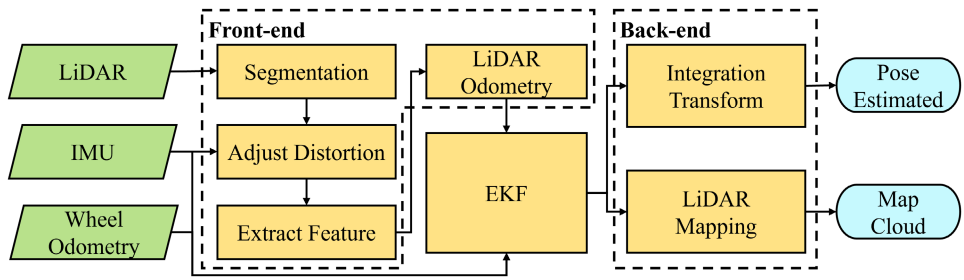
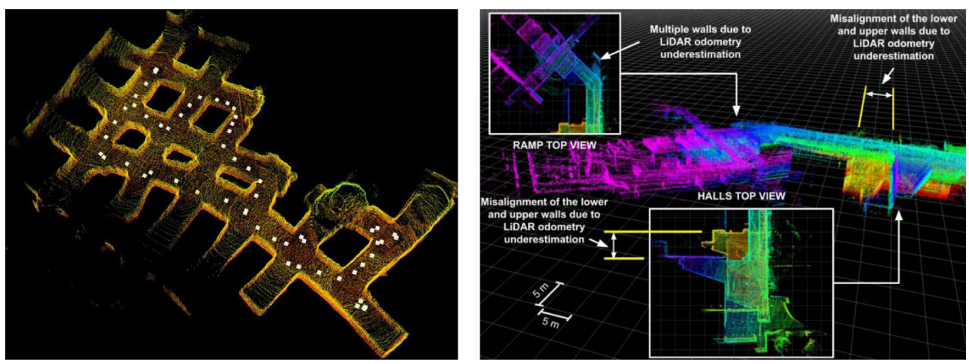


FIGURE 9 A representative framework of loosely coupled LI SLAM (Júnior et al., 2022).

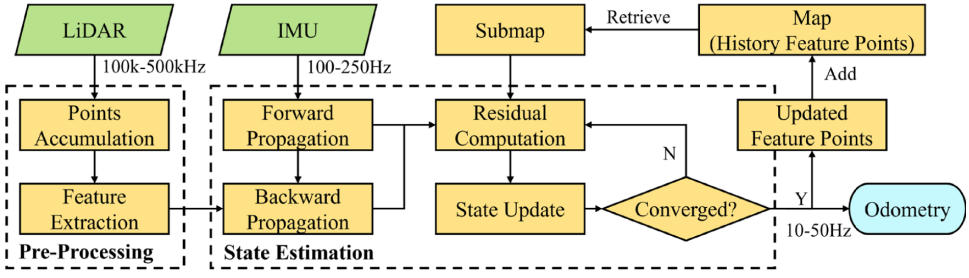


(a) Kim et al. [165] in Underground (b) EKF-LOAM [168] in Indoor

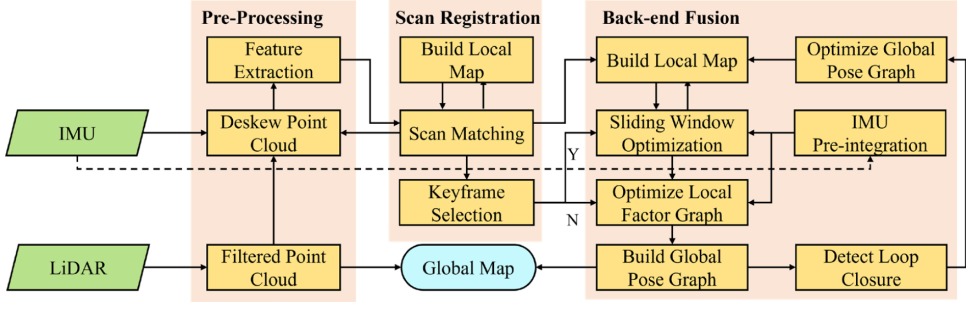
FIGURE 10 Mapping results of a loosely coupled LI SLAM.

be divided into filter-based methods and graph optimization methods. Filtering methods usually construct state models and measurement models based on the observations from various sensors such as LiDAR, IMU, etc., and carry out state estimation through the particle filter or Kalman filter. An early important work was proposed by Bry et al. (2012), which fused 2D LiDAR and IMU observations under the Gaussian particle filter framework and applied them to the Boston Dynamics Atlas robot. However, due to a large number of 3D LiDAR feature points, the time complexity of particle filtering is too high to run in real-time. Kalman filter and its variants have been introduced to replace particle filtering, such as the UKF (Holmes et al., 2008; Huang et al., 2013), EKF (Huang et al., 2008; Huang & Dissanayake, 2007), and iterated Kalman filter (Qin et al., 2020; Xu & Zhang, 2021). Since the UKF and EKF are susceptible to linearization errors, they are prone to produce false matches in scan matching when the initializations are not accurate enough, resulting in poor SLAM performance. The iterated Kalman filter-based methods perform error correction in each iteration, achieving the current SOTA performance.

LINS (Qin et al., 2020) designs an error-state iterated Kalman filter (ESIKF) to recursively correct the estimated state by generating new point correspondences in each iteration and reduces the number of features by fitting the ground plane to achieve real-time performance. Swarm-LIO (Zhu et al., 2023) implements a fully decentralized UAV swarm system state estimation method under the framework of ESIKF. Similarly, Fast-LIO (Xu & Zhang, 2021) (Figure 11a) realizes the tight coupling of LiDAR and IMU observations under the iterated extended Kalman filter (IEKF) framework. Compared with LINS, Fast-LIO proposes a new Kalman gain calculation formula, which makes the computational complexity of Kalman filtering depend on the dimension of the state vector rather than the feature dimension, thus greatly reducing the calculation load of the entire system. This also makes Fast-LIO the latest filtering-based baseline, and a number of variants have been developed. For example, Fast-LIO2 (Xu, Cai, et al., 2022) uses a direct method to replace the feature-based one in Fast-LIO for scan matching, and proposes an incremental k-d tree (kd-Tree) data structure for map maintenance, thereby improving the accuracy and speed. Faster-LIO (Bai et al., 2022) adopts sparse incremental voxels (iVox) to replace the tree structure for point cloud



(a) Filter-based Method

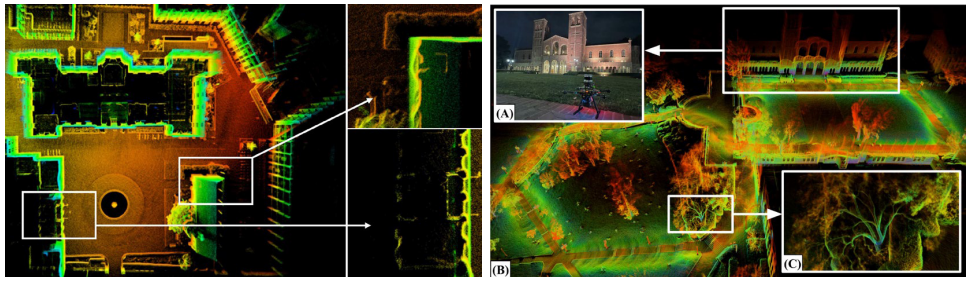


(b) Graph optimization-based Method

FIGURE 11 Two representative frameworks of tightly coupled LI SLAM. (a) and (b) are originally shown in Xu & Zhang (2021) and Li, Li, & Hanebeck (2021), respectively.

organization, further improving the computing efficiency of Fast-LIO2. Puma-LIO (Jiang & Shen, 2022) considers the principled uncertainty model in LiDAR-inertial odometry (LIO). AdaLIO (Lim et al., 2023) adds a surrounding environment checking module to adaptively adjust the parameters for scenes such as corridors, thus improving the performance of the algorithm in degraded scenes. This type of method achieves high accuracy and efficiency. However, since the loop closure constraint is not considered in the system, the problem of error accumulation may be relatively obvious in large-scale scenes. To address this problem, HD-LIO (Wang et al., 2023) (Figure 12a) and (Duan et al., 2023) add a LCD module and a factor graph optimization module after the IEKF module, but this inevitably increases the computational complexity of the system.

The graph optimization methods incorporate LiDAR and IMU factors into a unified graph structure (usually expressed as a factor graph or a pose graph), which is described as an MAP estimation problem and optimized using tools such as GTSAM, Ceres, and G2O. An early work, LIPS (Geneva et al., 2018), integrates plane constraints of LiDAR scans and IMU pre-integration constraints in the graph optimization. LIOM (Ye et al., 2019) is based on edge and plane features and uses a local sliding window for smoothing. It also presents a rotation constraint to refine the LiDAR pose. A more influential work is LIO-SAM (Shan et al., 2020), which combines keyframes and local sliding window strategies to improve the real-time performance of LIOM and fuses four factors in graph optimization, namely LiDAR odometry factor, IMU pre-integration factor, loop closure factor, and optional GPS factor. However, LIO-SAM relies on a high-frequency nine-axis IMU, which limits its popularization. As depicted in Figure 11b, LiLi-OM (Li, Li, & Hanebeck, 2021) explores the tight coupling between solid-state LiDAR and IMU. It designs a new feature extraction method based on the characteristics of solid-state LiDAR and uses strategies such as sliding windows, local factor graphs, and global factor graphs for optimization. LIO-vehicle (Xiao et al., 2022) is a SLAM method dedicated to vehicle trajectory estimation. It adds low-cost sensor (e.g., wheel speedometer and steering angle sensor) observations and vehicle motion constraints to the LIO-SAM framework to construct a more accurate factor graph. Similarly, GR-LOAM (Su et al., 2021) adds encoder constraints to improve the robustness of ground robots in complex terrain scenes. The scan matching of the above methods are all feature-based methods, in addition, there are some methods based on direct matching. As stated by Xu, Cai, et al. (2022), direct methods have better registration accuracy. Koide et al. (2021b) used a GPU-accelerated voxelized GICP (Koide et al., 2021b) to build a matching cost factor and fused local mapping and global mapping to refine the results of LIO. The architecture of D-LIOM (Wang et al., 2022) is similar to LiLi-OM, except that it directly registers the scans with probabilistic submaps to replace feature matching. As described in Figure 12b, DLIO (Chen et al., 2023) combines the ideas of direct method and continuous-time motion to achieve accurate distortion correction by considering the motion state estimation of each 3D point within a single frame. In general, graph optimization achieves high performance, especially in error accumulation suppression. However, these methods are computationally intensive, and strategies such as keyframes and local sliding windows are usually used to achieve real-time operation. Keyframes mean that many LiDAR scans must be discarded, which may cause information loss.



(a) HD-LIO [174] in Campus (b) DLIO [182] in Park

FIGURE 12 Mapping results of tightly coupled LI SLAM.

LiDAR-visual (LV) fusion

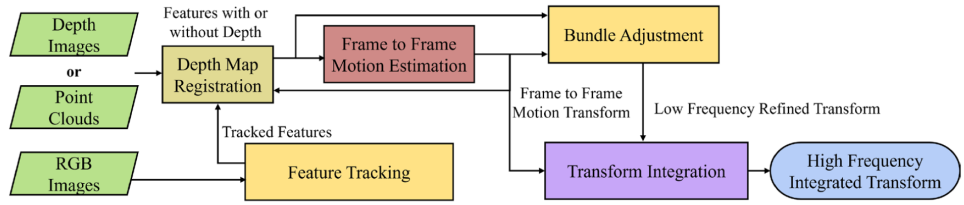
Visual SLAM has the advantages of rich texture information and small data volume but suffers from light sensitivity and scale drift. Moreover, maps built by visual SLAM methods are usually very sparse. In contrast, LiDAR SLAM directly acquires dense 3D information at the real scale with illumination invariance, but it lacks texture and the data is disorganized. It can be seen that there is a good complementary relationship between them. Therefore, LV fusion can effectively improve the accuracy and reliability of the SLAM system, and a dense 3D environment map with texture consistent with the real world can be obtained. Generally, LV fusion methods can be categorized into LiDAR-assisted visual SLAM methods and LV coupled SLAM ones. Table 5 gives a summary of these methods.

LiDAR-assisted visual SLAM

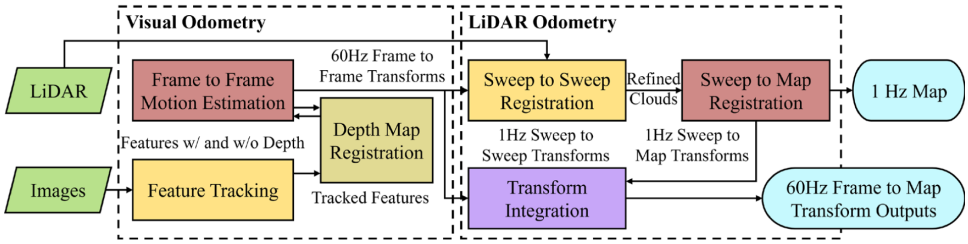
These methods utilize the LiDAR sensor to provide depth information for a visual SLAM system, so that it has an accurate scale factor, thereby improving the robustness of large-scale mapping applications (see Figure 13(a) for an example). It can be seen that the core of this type of method is still visual SLAM, which can be divided into

TABLE 5 A summary of LiDAR-visual fusion LiDAR SLAM. These methods are organized based on coupled types, method names, published years and code websites.

Type	Feature	Method	Year	Code
LiDAR-assisted visual SLAM	Feature method	DEMO (Zhang et al., 2017)	2014	×
		LIMO (Graeter et al., 2018)	2018	https://github.com/johannes-graeter/limo
		PL-LIMO (Huang et al., 2020)	2020	×
		CamVox (Zhu et al., 2021)	2021	https://github.com/ISEE-Technology/CamVox
		H-VLO (Aydemir et al., 2022)	2022	×
		RGBL (Sauerbeck et al., 2023)	2023	https://github.com/TUMFTM/ORB_SLAM3_RGBL
	Direct method	DVL-SLAM (Shin et al., 2020)	2018	http://github.com/irapkaist/dvl_slam
		SelfCompDVLO (Song et al., 2021)	2022	https://github.com/ZhenboSong/SelfCompDVLO-pytorch
		CR-LDSO (Yuan, Cheng, & Yang, 2023)	2023	×
LiDAR-visual coupled SLAM	Loose coupling	VLOAM (Zhang & Singh, 2015)	2015	×
		DV-LOAM (Wang, Liu, Wang, et al., 2021)	2021	https://github.com/kinggreat24/dv-loam
		VLO-SCVB (Cai et al., 2023)	2023	×
		SDV-LOAM (Yuan, Wang, et al., 2023)	2023	https://github.com/ZikangYuan/SDV-LOAM
	Tight coupling	TVLO (Seo & Chou, 2019)	2019	×
		ViLiVO (Xiang et al., 2019)	2019	×
		TVL-SLAM (Chou & Chou, 2022)	2021	×
		MMFC-TVL-SLAM (Shu & Luo, 2022)	2022	×

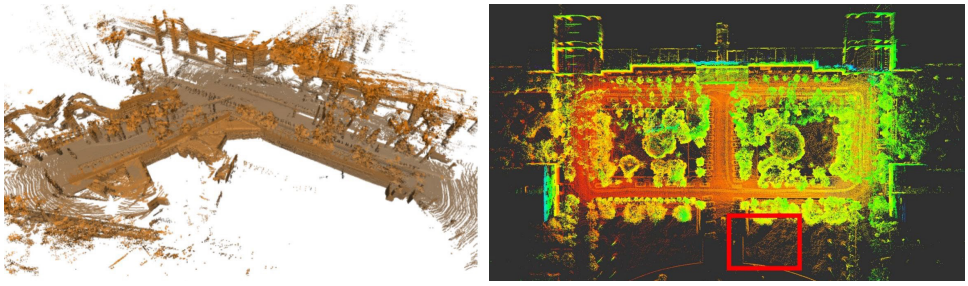


(a) LiDAR-assisted visual SLAM



(b) LiDAR-visual coupled SLAM

FIGURE 13 Two representative frameworks of LV fusion LiDAR SLAM. (a) and (b) are originally shown in Zhang et al. (2017) and Zhang & Singh (2015), respectively.



(a) PL-LIMO [192] in Urban (b) DV-LOAM [200] in Campus

FIGURE 14 Mapping results of LV fusion LiDAR SLAM.

feature-based methods and direct ones. Feature-based methods extract salient visual features from images and use LIDAR to provide depth information to them, which is then fed into a visual SLAM system for localization and mapping. Representative methods include algorithms based on point features such as DEMO (Zhang et al., 2017), LIMO (Graeter et al., 2018), and CamVox (Zhu et al., 2021), and the ones that fuse point and line features such as PL-LIMO (Huang et al., 2020) (Figure 14a). However, due to the sparsity of LIDAR point clouds, only a very small number of visual features can be associated with effective depth information and interpolation errors will be introduced during the association. To alleviate this problem, RGBL (Sauerbeck et al., 2023) complements the LiDAR depth map to obtain an RGBD image, which is then fed into the ORB-SLAM3 (Campos et al., 2021) system; H-VLO (Aydemir et al., 2022) uses a deep neural network to perform LiDAR depth completion and monocular depth estimation at the same time and provides scale correction for the visual SLAM system through the registration of LiDAR depth map and monocular depth map. However, the erroneous depth information in depth completion and depth estimation will have an impact on the robustness of the SLAM system.

Direct methods can effectively deal with the problems of insufficient depth features and interpolation errors. Instead of detecting features, they project LiDAR point clouds onto images and then feed the pixels with depth information into the direct sparse odometry (DSO) (Engel et al., 2018) system for tracking. Since DSO

directly minimizes a photometric error between images, any region containing pixel gradients in the image can be used as the object tracked by DSO, such as white wall edges, intensity change areas, etc. Therefore, DSO is more robust to weak/texture-less scenes. DVL-SLAM (Shin et al., 2020) is the first to introduce DSO to realize LV fusion; CR-LDSO (Yuan, Cheng, & Yang, 2023) introduces cloud reuse strategy in DSO to solve the problem of insufficient tracking pixels; SelfCompDVLO (Song et al., 2021) performs a LiDAR depth completion step and replaces the photometric-metric error in DSO with feature-metric error. However, since the direct methods achieve tracking based on pixel intensity information, they are prone to fail in situations where the illumination changes dramatically.

LV coupled SLAM

Different from LiDAR-assisted visual SLAM, this type of method combines the LO factor and the visual odometry (VO) factor to improve the accuracy and robustness of the system (see Figure 13b for an example). Again, these methods contain both loosely and tightly coupled approaches. Loose coupling usually uses a cascade strategy to combine the LO and VO modules. A pioneering work is VLOAM (Zhang & Singh, 2015), which uses a high-speed VO module (running at 60 Hz) to obtain low-fidelity poses, and then refines the motion estimates and corrects distortions in LiDAR scans through a low-speed (running at 1 Hz) LO module. It achieved the best accuracy at that time on KITTI. VLO-SCVB (Cai et al., 2023), SDV-LOAM (Yuan, Wang, et al., 2023), DV-LOAM (Wang, Liu, Wang, et al., 2021) (Figure 14b) are variant methods under the framework of VLOAM. To overcome the above-mentioned interpolation error problem, VLO-SCVB first projects the LiDAR points onto images then selects key points in the projected points, and fuses the results of ORB-SLAM2 (Mur-Artal & Tardós, 2017) and LOAM in a loose coupling way. To address the problems of insufficient depth features and interpolation errors, SDV-LOAM combines the advantages of feature-based methods and direct ones by using photometric consistency for initial tracking and then minimizing the reprojection errors of features for further refinement. It is a semi-direct method. DV-LOAM replaces the feature-based approach in VLOAM with a direct one. Although loose coupling has the advantages of simple system structure and high local accuracy, its robustness to large and complex scenes is not high enough.

Tight coupling generally formulates the state estimation factors of LiDAR and vision observations as a joint optimization problem (Cao et al., 2023; Chou & Chou, 2022; Seo & Chou, 2019). TVLO (Seo & Chou, 2019) simultaneously constructs a LiDAR voxel map and a visual sparse map and integrates the residuals of LO and VO into the same objective function for optimization to achieve accurate pose tracking. Similarly, TVL-SLAM (Chou & Chou, 2022) and MMFC-TVL-SLAM (Shu & Luo, 2022) incorporate all LiDAR and visual observations into the factor graph of the SLAM back end (including visual factors, LiDAR factors, compose factor, and LV factors), thereby achieving joint optimization of LiDAR and visual residuals. ViLiVO (Xiang et al., 2019) proposes a concept of virtual LiDAR based on visual images. Specifically, the free space in an image is extracted through a semantic segmentation model, and then the contours of the free space are discretized to generate the scan lines of the virtual LiDAR. Tight coupling usually has higher global accuracy and robustness, but the system structure is more complex and the computational complexity is large.

LiDAR-inertial-visual (LIV) fusion

LiDAR can directly obtain dense 3D information and is invariant to illumination changes; IMU has a high output frequency, which can effectively improve the accuracy in featureless scenes such as corridors; visual images contain rich texture information, which makes feature extraction and LCD reliable. Therefore, integrating the advantageous information of these different modalities can improve the system's ability to deal with complex scenarios. Like LI fusion, loose coupling and tight coupling are the two most important coupling methods. Table 6 presents a summary of LIV fusion LiDAR SLAM.

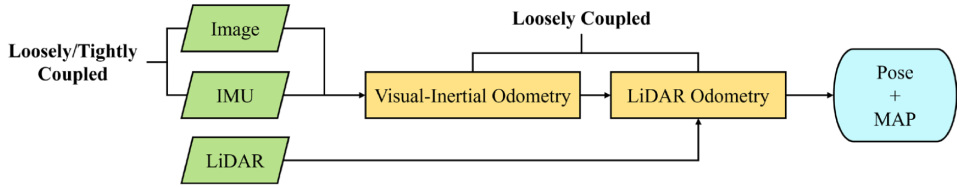
TABLE 6 A summary of LiDAR-inertial-visual LiDAR SLAM. These methods are organized based on coupled types, method names, published years and code websites. T1, T2 and T3 denote that the quantity of tightly coupled modules is one, two and three, respectively.

Type	Feature	Method	Year	Code
Loosely coupled LIV SLAM	T1	VI-LOAM (Zhang & Singh, 2018)	2018	×
		VIL-SLAM (Shao et al., 2019)	2019	×
		VILO (Wang, Zhang, et al., 2019)	2019	×
		VIO-LOAM (Khattak et al., 2020)	2020	×
	T2	MetroLoc (Wang, Song, Zhang, et al., 2021)	2021	×
		R3LIVE (Lin & Zhang, 2022a)	2022	https://github.com/hku-mars/r3live
		R3LIVE++ (Lin & Zhang, 2022b)	2022	https://github.com/hku-mars/r3live
Tightly coupled LIV SLAM	T3 Filter-based	LIC-Fusion (Zuo et al., 2019)	2019	×
		LIC-Fusion2 (Zuo et al., 2020)	2020	×
		R2LIVE (Lin et al., 2021)	2021	https://github.com/hku-mars/r2live
		FAST-LIVO (Zheng et al., 2022)	2022	https://github.com/hku-mars/FAST-LIVO
	T3 Factor graph	LVI-SAM (Shan et al., 2021)	2021	https://git.io/lvi-sam
		VILENS-LVI (Wisth et al., 2021)	2021	×
		LVIO-FUSION (Jia et al., 2021)	2021	×
		Super odometry (Zhao et al., 2021)	2021	https://sites.google.com/view/superodometry
		VILENS (Wisth et al., 2022)	2022	×
		CLIC (Lv et al., 2023)	2023	https://github.com/APRIL-ZJU/clic

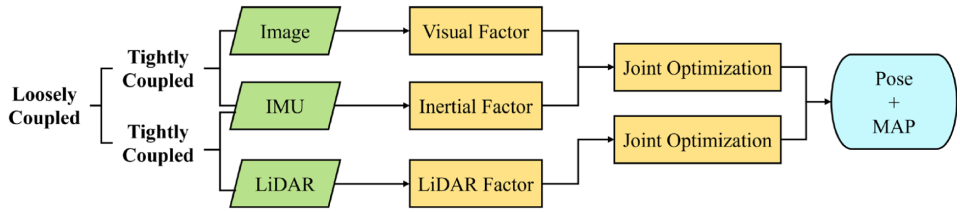
Loosely coupled LIV SLAM

Loose coupling means that the state estimation factors of LiDAR, visual, and inertial are not jointly optimized. It mainly contains two cases, whose system architectures are shown in Figure 15a,b, respectively. The first case is that visual and inertial are loosely or tightly coupled to obtain the visual-inertial odometry (VIO) motion estimates and provide them as initializations to a LO module for refinement (An example is provided in Figures 16a and 17a). In this case, only the visual and inertial factors may have been jointly optimized, with the most representative work being VI-LOAM (Zhang & Singh, 2018). However, VI-LOAM performs frame-to-frame motion estimation, which is highly dependent on the previous frame and difficult to maintain global consistency. To improve global consistency, VILO (Wang, Zhang, et al., 2019) adds LCD and back-end optimization modules in the VI-LOAM; VIL-SLAM (Shao et al., 2019) implements a stereo camera version of VILO; VIO-LOAM (Khattak et al., 2020) introduces thermal image information into the VIO module of VI-LOAM to improve its robustness in degraded scenes (such as tunnels, mines and other low-light scenes). Although this coupling strategy is very simple to implement and has high computational efficiency, it does not take full advantage of various types of observation factors.

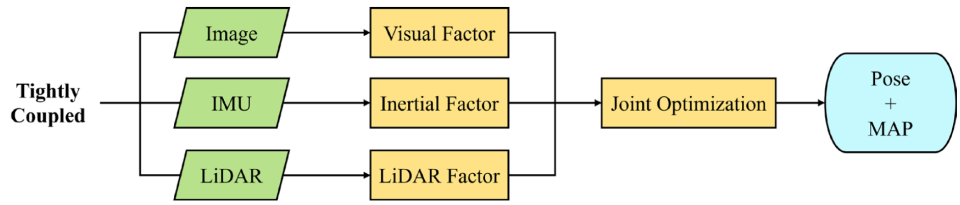
The system architecture of the second case is shown in Figure 15(b), which contains two tightly coupled modules, that is, the VIO and LIO modules. However, these two modules are fused in a loosely coupled way to obtain the final motion estimation results (see Figure 16b for an example). Although this coupling strategy is more 'tight' than the previous one and is referred to as tight coupling in some papers (Lin & Zhang, 2022a, 2022b), the state estimation factors of LiDAR, vision, and inertial are not jointly optimized under the same objective function. Therefore, it is essentially a kind of loose coupling, with representative approaches such as MetroLoc (Wang, Song, Zhang, et al., 2021), R3LIVE (Lin & Zhang, 2022a), and R3LIVE++ (Lin & Zhang, 2022b) (Figure 17b). MetroLoc constructs an IMU-centric factor graph estimator, where its LIO is based on line and planar features, and its VIO is built on the Vins-mono (Qin et al., 2018). The authors claim that the combination of tight and loose coupling is able to merge the advantageous properties of both coupling strategies. R3LIVE and R3LIVE++ are filtering-based



(a) Loosely coupled LIV SLAM with a single tightly coupled module



(b) Loosely coupled LIV SLAM with two tightly coupled module



(c) Tightly coupled LIV SLAM

FIGURE 15 Three LIV system architectures.

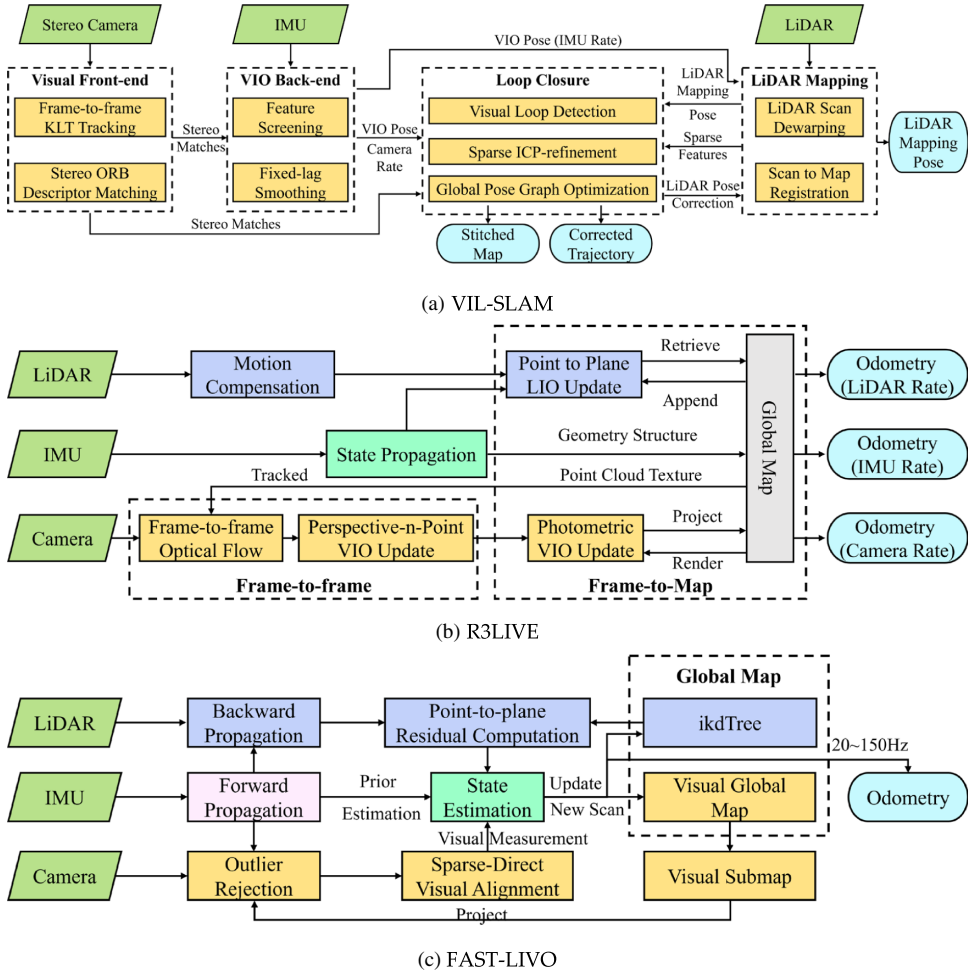
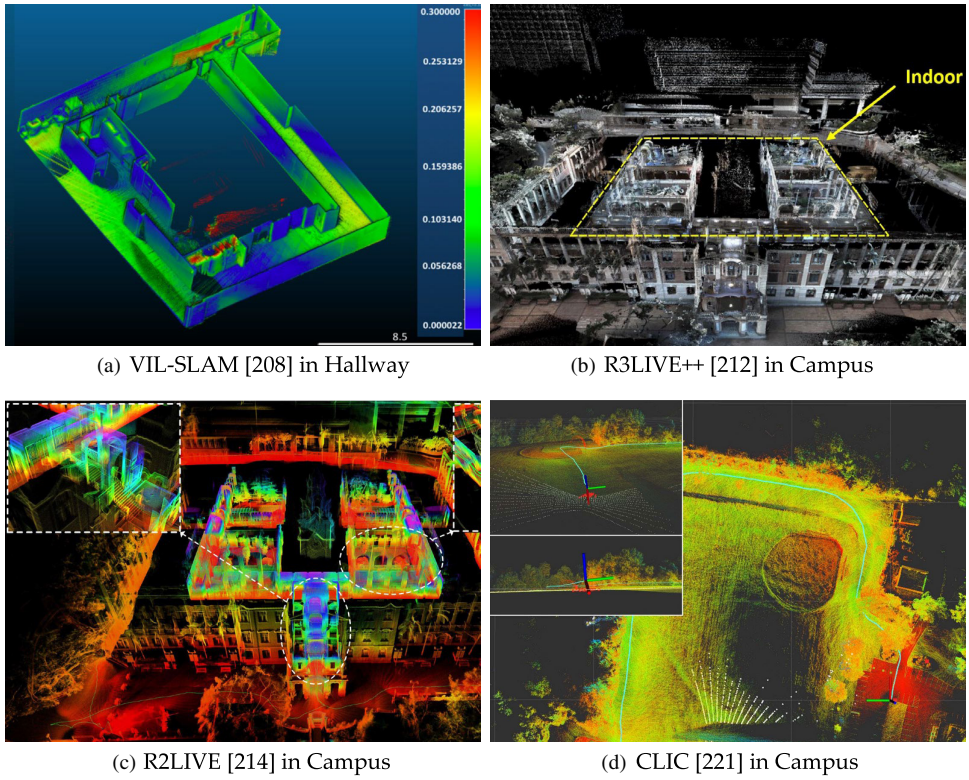


FIGURE 16 Three LIV LiDAR SLAM methods. (a–c) are originally shown in Shao et al. (2019), Lin & Zhang (2022a) and Zheng et al. (2022), respectively.

methods, whose LIO modules are built on the FastLIO2. Their VIO modules use a semi-direct method to combine reprojection errors and photometric errors and apply the ESIKF (Qin et al., 2020; Xu & Zhang, 2021) for state estimation. Generally, this type of coupling has higher accuracy and robustness than the previous one.

Tightly coupled LIV SLAM

The system architecture of tight coupling is shown in Figure 15c. Depending on the optimizer, tightly coupled methods can be also classified into filtering-based methods and graph optimization-based ones. A typical early filtering method is LIC-Fusion (Zuo et al., 2019), which incorporates the three types of observations using the multi-state constraint Kalman filter (MSCKF) (Mourikis & Roumeliotis, 2007) framework. Subsequently, in order to improve the efficiency, LIC-Fusion2 (Zuo et al., 2020) introduces a sliding-window plane-feature tracking module under the framework of LIC-Fusion to process LiDAR point clouds. Similarly, R2LIVE (Lin et al., 2021) (Figure 17c) replaces the MSCKF in LIC-Fusion with the ESIKF and adds a factor graph optimization module to further refine the filtering results. As can be seen, R2LIVE is a hybrid method of filtering and factor graph optimization. To reduce the amount of computation, FAST-LIVO (Zheng et al., 2022) (Figure 16c) directly registers original LiDAR



(a) VIL-SLAM [208] in Hallway (b) R3LIVE++ [212] in Campus

(c) R2LIVE [214] in Campus (d) CLIC [221] in Campus

FIGURE 17 Mapping results of LIV SLAM. (a) and (b) are loosely coupled LIV SLAM, while (c) and (d) are tightly coupled ones.

sample points in the LIO module and uses a semi-direct method for visual tracking in the VIO module, avoiding the feature extraction and description process in the whole system. Like R2LIVE, it uses ESIKF for joint filtering.

LVI-SAM (Shan et al., 2021) is a typical factor graph optimization method, which contains a VIO module based on Vins-mono and a LIO module based on LIO-SAM. Among them, LIO provides initializations for VIO, and VIO provides motion priors for the scan matching of LIO. LVI-SAM uses the iSAM2 to jointly optimize the IMU constraint, VO constraint, LO constraint, and loop closure constraint in the factor graph. VILENS-LVI (Wisth et al., 2021) extracts plane and linear features as registration primitives for LO constraints in the factor graph; VILENS (Wisth et al., 2022) introduces leg odometry constraints on the basis of VILENS-LVI to improve the localization accuracy of legged robots. Lvio-fusion (Jia et al., 2021) uses a stereo camera to replace the monocular camera in LVI-SAM and introduces a GPS constraint factor to eliminate accumulated drifts, which has good robustness. CLIC (Lv et al., 2023) (Figure 17d) expresses the trajectory based on the B-spline and derives a continuous-time fixed-lag smoother to achieve continuous-time LIV SLAM. Super odometry (Zhao et al., 2021) provides a new idea that is centred on IMU odometry. It associates VIO and LIO modules in a combination of tight and loose coupling and uses a coarse-to-fine approach for motion estimation.

Summary

This section classifies the multi-sensor fusion SLAM into three categories, namely LI, LV, and LIV. It summarizes the methods in each category and describes the representative algorithms in detail. Key observations are summarized as follows:

- LiDAR sensors directly acquire dense 3D information of the surrounding environment, with the advantages of high measurement accuracy, illumination invariance, and no scale drift; IMU can measure observations such as angular velocity and acceleration during the motion process, and its output frequency and local accuracy are very high; visual images are direct 2D reconstructions of the real world, containing rich texture information and strong comprehensibility. These sensors have a very good complementary relationship. Through the fusion technology to synthesize the advantages of these different sensors, the system can improve its ability to deal with complex scenarios. Multi-sensor fusion has shown higher accuracy and robustness in real applications, which is one of the development trends of SLAM research.
- Loose coupling and tight coupling are the two main means to realize multi-sensor information fusion. Loose coupling can separate different modules of the SLAM system, thereby reducing the computational load, and has the advantages of running efficiently, easy to implement, and easy to expand. However, loose coupling cannot maximize the advantages of each sensor. In contrast, tight coupling optimizes various types of observation factors such as LiDAR, IMU, visual, loop closure, etc. in an objective function, which exhibits better performance in complex scenarios with higher accuracy and robustness, and thus is more popular in multi-sensor fusion.
- Tight coupling can be divided into filter-based methods and graph optimization methods. The filtering methods construct state models and measurement models based on sensor observations to realize fusion, which can achieve good real-time performance while fully utilizing each LiDAR scanning frame. However, filtering methods do not consider loop closure constraints, and the error accumulation is relatively obvious in large-scale scenes. Graph optimization usually expresses the observation factors as factor graphs and describes them as a MAP problem for optimization. Graph optimization has better performance in error accumulation suppression. However, it is computationally intensive and requires the use of keyframes to obtain real-time operation. Keyframes mean that many LiDAR scans have to be discarded and there may be information loss.
- Theoretically, the more types of sensors, the richer the information that the SLAM system can utilize, and the higher the accuracy and robustness it can achieve. However, more sensor types also mean that the SLAM system is more complex and less scalable. How to characterize the quality of different sensor data? How to adaptively detect erroneous observations in the fusion system and adjust the weights of different types of observations? And which fusion mechanism to use? These are the core issues that need to be addressed.

BENCHMARK

Datasets

Numerous teams have curated datasets for perception, localization, mapping, and tracking evaluation to advance autonomous driving. Table 7 compiles several notable LiDAR SLAM datasets with their published year, platform, type, sensor, and website. To assess SLAM performance across various challenging conditions, mainstream datasets frequently employ autonomous vehicles (Barnes et al., 2020; Caesar et al., 2020; Chang et al., 2019; Chen et al., 2018; Choi et al., 2018; Geiger et al., 2012, 2013; Geyer et al., 2020; Huang et al., 2018; Jeong et al., 2018, 2019; Liao et al., 2023; Nguyen et al., 2022; Pandey et al., 2011; Pitropov et al., 2021; Ramanishka et al., 2018) for data collection in expansive urban (Barnes et al., 2020; Caesar et al., 2020; Chang et al., 2019; Chen et al., 2018; Choi et al., 2018; Geiger et al., 2012, 2013; Geyer et al., 2020; Huang et al., 2018; Jeong et al., 2018, 2019; Liao et al., 2023; Pandey et al., 2011; Pitropov et al., 2021; Ramanishka et al., 2018) and campus environments (Choi et al., 2018; Nguyen et al., 2022; Pandey et al., 2011). Several sequences (Carlevaris-Bianco et al., 2016; Geiger et al., 2012, 2013; Liao et al., 2023; Pandey et al., 2011) incorporate closed trajectory to investigate the impact of loop detection modules on the overall algorithm. For indoor applications, some datasets employ low-speed robots (Carlevaris-Bianco et al., 2016; Lee et al., 2021) or

TABLE 7 Summary of notable public datasets.

Year	Dataset	Platform	Type	LiDAR	IMU	Camera	Website
2011	Ford Campus (Pandey et al., 2011)	Car	C	Velodyne HDL-64E Riegl LMS×2	100	Ladybug3 Camera	http://robots.engin.umich.edu/SoftwareData/Ford
2012	KITTI (Geiger et al., 2012, 2013)	Car	U, R	Velodyne HDL-64E	10	PointGrey (Grey) ×2 PointGrey (Colour) ×2	https://www.cvlibs.net/datasets/kitti
2016	NCLT (Carlevaris-Bianco et al., 2016)	Wheeled robot	C	Velodyne HDL-32E×2	100	Ladybug3 Camera	http://robots.engin.umich.edu/SoftwareData/NCLT
2016	SYNTHIA (Ros et al., 2016)	Synthetic	U	n.a.	n.a.	n.a.	http://adas.cvc.uab.es/synthia
2018	Complex Urban (Jeong et al., 2018)	Car	U	Velodyne VLP-16×2 SICK-LMS-511×2	100	n.a.	http://irap.kaist.ac.kr/dataset
2018	KAIST Multi-spectral (Choi et al., 2018)	Car	U, C	Velodyne HDL -32E	25	PointGrey (Colour)×2 Thermal camera	https://github.com/unizard/kaist-allday-dataset
2018	Apolloscape (Huang et al., 2018)	Car	U	VUX-1HA×2	n.a.	VMX-CS6	https://apolloscape.auto
2018	LiVi-Set (Chen et al., 2018)	Car	U	Velodyne HDL-32E Velodyne VLP-16	n.a.	Dashboard camera	https://github.com/driving-behavior/DBNet
2018	HDD (Ramanishka et al., 2018)	Car	U	Velodyne HDL-64E	n.a.	Grasshopper×3	https://usa.honda-ri.com/HDD
2019	Argoverse (Chang et al., 2019)	Car	U	Velodyne VLP-3 2×2	n.a.	Stereo Camera×2 Ring Camera×7	https://github.com/argoverse/argoverse-api
2019	KAIST Urban (Jeong et al., 2019)	Car	U	Velodyne VLP-16×2 SICK LMS-511×2	200	FLIR×2	http://irap.kaist.ac.kr/dataset
2020	A2D2 (Geyer et al., 2020)	Car	U	Velodyne VLP-16×5	n.a.	Sekonix×6	http://www.a2d2.audi

TABLE 7 (Continued)

Year	Dataset	Platform	Type	LiDAR	IMU	Camera	Website
2020	Oxford Radar RobotCar (Barnes et al., 2020)	Car	U	Velodyne HDL-32E×2 SICK LMS-151×2	50	Bumblebee Grasshopper2×3	https://oxford-robotics-institute.github.io/radar-robotcar-dataset
2020	nuScences (Caesar et al., 2020)	Car	U	32-beam LiDAR	1000	CMOS×6	https://nuscenes.org
2020	Newer College (Ramezani et al., 2020; Zhang et al., 2021)	Hand	C	Ouster-OS1-64	400	Intel Realsense	https://ori.ox.ac.uk/datasets/newer-college-dataset
2021	Navers Lab Indoor (Lee et al., 2021)	Wheeled robot	I	Velodyne VLP-16×2	n.a.	Basler×6	https://naverlabs.com/datasets
2021	CADC (Pitropov et al., 2021)	Car	U	Velodyne VLP-32C	200	Ximea×8	https://github.com/mpitropov/cadc_devkit
2022	S3LI (Giubilato et al., 2022)	Hand	P	Blickfeld Cube-1	400	AVT Mako×2	https://rmc.dlr.de/s3li_dataset
2022	NTU VIRAL (Nguyen et al., 2022)	UAV	C, I	Ouster-16×2	385	IDS (Grey) ×2	https://ntu-aris.github.io/ntu_viral_dataset
2023	WHU-Helmet (Li, Wu, et al., 2023)	Helmet	I, U, F	Livox	n.a.	n.a.	https://github.com/kafeiyin00/WHU-HelmetDataset
2023	KITTI-360 (Liao et al., 2023)	Car	U	Velodyne HDL-64E SICK LMS 200	10	Fisheye Camera×2 Stereo Camera	http://www.cvlibs.net/datasets/kitti-360

Note: C, campus; F, forest; I, indoor; P, planetary analogue; R, rural; U, urban.

handheld devices (Ramezani et al., 2020). Giubilato et al. (2022) and Li, Wu, et al. (2023) additionally introduce a specialized helmet dataset encompassing indoor, urban, and forest environments, while Ros et al. (2016) presents a synthetic dataset. These datasets encompass a range of LiDAR sensor densities, including 16-line (Carlevaris-Bianco et al., 2016; Jeong et al., 2018; Lee et al., 2021), 32-line (Caesar et al., 2020; Chang et al., 2019; Pitropov et al., 2021), 64-line (Geiger et al., 2012, 2013; Ramezani et al., 2020; Zhang et al., 2021), solid-state (Li, Wu, et al., 2023), and dual LiDARs (Chen et al., 2018; Jeong et al., 2019; Nguyen et al., 2022). To facilitate comparisons with pure vision, LiDAR-vision, and IMU-LiDAR solutions, diverse camera (Barnes et al., 2020; Caesar et al., 2020; Chang et al., 2019) and IMU (Choi et al., 2018; Giubilato et al., 2022; Nguyen et al., 2022) set-ups were integrated into the mobile platform.

Evaluation metrics

Numerous metrics are available for assessing SLAM performance. Here, the top three commonly employed metrics in LiDAR-based SLAM are illustrated, accompanied by visual representations and mathematical expressions:

Relative pose error (RPE)

As shown in Figure 18(a), RPE (Sturm et al., 2012) gauges the precision of relative poses, also known as odometry, within a SLAM trajectory. RPE is insensitive to the specific timing of estimation errors. It encompasses both translational and rotational errors, playing a pivotal role in ensuring the accuracy of map construction, updating, and sensor motion determination in the environment. As Geiger et al. (2012) suggest, RPE is calculated as:

$$\begin{aligned} RPE_{rot}(\mathcal{F}) &= \frac{1}{\mathcal{F}} \sum_{i,j \in \mathcal{F}} \angle \left[\hat{T}_j \hat{T}_i^{-1} T_i T_j^{-1} \right] \\ RPE_{tra}(\mathcal{F}) &= \frac{1}{\mathcal{F}} \sum_{i,j \in \mathcal{F}} \left\| \hat{T}_j \hat{T}_i^{-1} T_i T_j^{-1} \right\|_2 \end{aligned} \tag{3}$$

where RPE_{rot} and RPE_{tra} demote the relative rotation error (RRE) and relative translation error (RTE), respectively. $\mathcal{F}$ is a set of frames (i, j). T and $\hat{T}$ denote the estimated and true poses, respectively. $\angle[\cdot]$ is the rotation angle.

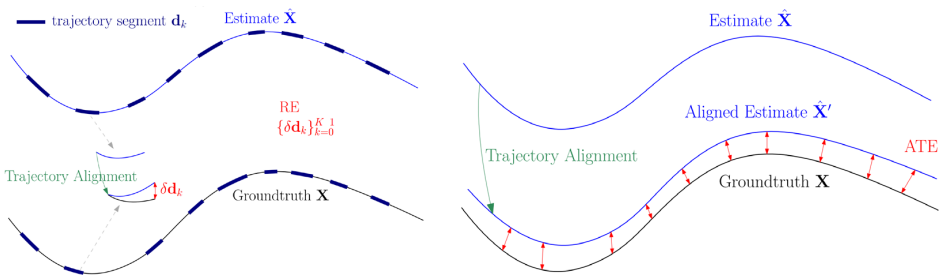


FIGURE 18 RPE and ATE. (a) and (b) are originally shown in Zhang & Scaramuzza (2018).

Absolute trajectory error (ATE)

Figure 18(b) illustrates that ATE (Zhang & Scaramuzza, 2018) quantifies absolute position and attitude errors relative to ground truth, while relative pose error computes average errors within sub-trajectories, potentially underestimating loop closing effects. Nonetheless, ATE facilitates a holistic comparison of entire trajectories, considering absolute error assessments modified by loop closing. As Zhang & Scaramuzza (2018) suggest, ATE is calculated as:

$$\begin{aligned} ATE_{rot} &= \left(\frac{1}{N} \sum_{i=0}^{N-1} \left\| \angle \left(R_i \widehat{R}_i' \right) \right\|^2 \right)^{\frac{1}{2}} \\ ATE_{tra} &= \left(\frac{1}{N} \sum_{i=0}^{N-1} \left\| \left(p_i - \angle \left(R_i \widehat{R}_i' \right) \widehat{p}_i' \right) \right\|^2 \right)^{\frac{1}{2}} \end{aligned} \tag{4}$$

where ATE_{rot} and ATE_{tra} demote the rotation and translation component of ATE, respectively. R and $\widehat{R}'$ are true and aligned rotation, respectively. p is the position of the system. i means the scan index and N is the scan quantity.

Running frequency

The running time in LiDAR SLAM is a critical factor that impacts mapping efficiency and overall robot product cost. Handcrafted methods involve time-consuming processes like segmentation, feature extraction, inter-frame matching, and mapping optimization. Learning-based methods require training and inference, further adding to the computational demands.

Results

To help researchers grasp the performance of contemporary LiDAR SLAM methods and gauge the current state of this field, comprehensive experimental evaluations on public datasets from various sources are compiled. The relative translational and rotational errors on KITTI odometry (Geiger et al., 2012) are summarized in Tables 8 and 9, respectively. These statistical results lead to the following conclusions:

- The emergence of numerous learning-based LiDAR SLAM methods in recent years underscores the growing interest in this field within computer vision. Researchers are actively exploring diverse solutions to address autonomous driving challenges.
- Statistical results reveal that among current LiDAR SLAM solutions, traditional methods exhibit superior translational and rotational accuracy. This superiority may be attributed to their comprehensive modules encompassing odometry, optimization, and mapping, whereas learning-based methods primarily emphasize front-end odometry, resulting in slightly reduced accuracy.
- Regarding position accuracy, MULLS-SLAM (Pan et al., 2021), IMLS-SLAM (Deschaud, 2018), and CT-ICP (Dellenbach et al., 2022) rank among the top three in three sequences, whereas MULLS-LO (Pan et al., 2021) and SDV-LOAM achieve top-three rankings in five and eight ones, respectively. When considering the average of 11 sequences, SDV-LOAM (Yuan, Wang, et al., 2023), MULLS-SLAM (Pan et al., 2021), and ELO (Zheng & Zhu, 2021) hold the first (0.47%), second (0.49%), and third places (0.50%), respectively.
- We also tally the top three angular accuracy rankings, with MULLS-SLAM (Pan et al., 2021) outperforming all others. Specifically, MULLS-LO (Pan et al., 2021), ELO (Zheng & Zhu, 2021), and reference (Koide et al., 2021a) achieve the top three on 11, six, and three sequences, respectively. MULLS-SLAM (Pan et al., 2021) secures the first position (0.13°) in average accuracy, followed by MULLS-LO (Pan et al., 2021) in the second (0.16°) and ELO

TABLE 8 Relative translational error (%) on KITTI odometry (Geiger et al., 2012).

Methods	Year	LP	00* U	01 U	02* U	03 U	04 U	05* U	06* U	07* U	08* U	09* U	10 U	00–10 mean
Handcrafted methods														
ICP-po2po (Besl & McKay, 1992)	1992		6.88	11.21	8.21	11.07	6.64	3.97	1.95	5.17	10.04	6.93	8.91	7.37
ICP-po2pl (Besl & McKay, 1992)			3.80	13.53	9.00	2.72	2.96	2.29	1.77	1.55	4.42	3.95	6.13	4.72
GICP (Segal et al., 2009)	2009		1.29	4.39	2.53	1.68	3.76	1.02	0.92	0.64	1.58	1.97	1.31	1.92
CLS (Velas et al., 2016)	2016		2.11	4.22	2.29	1.63	1.59	1.98	0.92	1.04	2.14	1.95	3.46	2.12
LOAM (Ji & Singh, 2017)	2017		0.78	1.43	0.92	0.86	0.71	0.57	0.65	0.63	1.12	0.77	0.79	0.84
DEMO (Zhang et al., 2017)			1.05	1.87	0.93	0.99	1.23	1.04	0.96	1.16	1.24	1.17	1.14	1.16
IMLS-SLAM (Deschaud, 2018)	2018		0.50	0.82	0.53	0.68	0.33	0.32	0.33	0.33	0.80	0.55	0.53	0.52
LeGO-LOAM (Shan & Englot, 2018)		✓	1.34	3.05	5.00	1.55	1.39	0.88	1.29	0.65	1.27	1.76	1.97	1.75
TVLO (Seo & Chou, 2019)	2019	✓	0.66	0.84	0.79	0.65	0.74	0.49	0.40	0.47	1.02	0.76	0.60	0.74
HDL-Graph-SLAM (Koide et al., 2019)		✓	1.29	58.40	8.09	1.54	1.83	1.44	0.54	1.85	2.05	2.00	3.06	7.46
FALO (García Daza et al., 2020)	2020		1.28	2.36	1.15	0.93	0.98	0.45	–	0.44	–	0.64	0.83	1.00
Huang et al. (Huang et al., 2020)		✓	0.99	1.87	1.38	0.65	0.42	0.72	0.61	0.56	1.27	1.06	0.83	0.94
LiTAMIN (Yokozuka et al., 2020)			0.91	11.3	1.30	1.17	18.7	0.75	0.59	0.48	1.04	0.99	3.78	1.60
LiTAMIN (Yokozuka et al., 2020)		✓	0.95	11.3	1.25	1.17	18.7	0.70	0.63	0.45	1.03	1.06	3.78	1.59

TABLE 8 (Continued)

Methods	Year	LP	00* U	01 U	02* U	03 U	04 U	05* U	06* U	07* U	08* U	09* U	10 U	00–10 mean
DV-LOAM (Wang, Liu, Wang, et al., 2021)	2021	✓	0.65	1.74	0.97	0.85	0.39	0.54	0.65	0.51	0.89	0.73	0.87	0.80
ELO (Zheng & Zhu, 2021)			0.54	0.61	0.54	0.65	0.32	0.33	0.30	0.31	0.79	0.48	0.59	0.50
F-LOAM (Wang, Wang, Chen, & Xie, 2021)			0.92	2.80	1.56	1.09	1.43	0.79	0.72	0.54	1.11	1.28	1.77	1.27
ISC-LOAM (Wang, Wang, & Xie, 2021)		✓	1.02	2.92	1.67	1.15	1.50	0.81	0.76	0.56	1.20	1.40	1.87	1.35
GLIM (Koide et al., 2021a)		✓	0.58	0.61	0.60	0.69	0.44	0.38	0.34	0.37	1.51	0.68	0.74	0.74
LiTAMIN2 (Yokozuka et al., 2021)			0.75	1.88	0.99	0.84	0.90	0.74	0.45	0.55	1.25	0.74	1.36	0.95
LiTAMIN2 (Yokozuka et al., 2021)		✓	0.70	2.10	0.98	0.96	1.05	0.45	0.59	0.44	0.95	0.69	0.80	0.85
MULLS-LO (Pan et al., 2021)			0.51	0.62	0.55	0.61	0.35	0.28	0.24	0.29	0.80	0.49	0.61	0.49
MULLS-SLAM (Pan et al., 2021)		✓	0.54	0.62	0.69	0.61	0.35	0.29	0.29	0.27	0.83	0.51	0.61	0.52
S4-SLAM (Zhou, He, et al., 2021)		✓	0.62	1.11	1.63	0.82	0.95	0.50	0.65	0.60	1.33	1.05	0.96	0.92
Vizzo et al. (Vizzo et al., 2021)			1.55	1.46	3.38	1.86	1.60	1.63	1.20	0.88	0.72	1.44	1.51	1.38
Aydemir et al. (Aydemir et al., 2022)	2022		0.99	1.87	1.38	0.65	0.42	0.72	0.61	0.56	1.27	1.06	0.83	0.94
CT-ICP (Dellenbach et al., 2022)		✓	0.49	0.76	0.52	0.72	0.39	0.25	0.27	0.31	0.81	0.49	0.48	0.53
E-LOAM (Guo et al., 2023)		✓	1.17	2.92	2.36	1.15	1.38	0.82	1.34	1.21	1.63	1.36	1.84	1.56
NDT-LOAM (Chen, Ma, et al., 2022)			0.79	1.46	1.09	0.65	0.31	0.54	0.56	0.27	1.04	0.74	1.12	0.90
T-LOAM (Zhou, Guo, et al., 2022)			0.98	2.09	1.01	1.10	0.68	0.55	0.56	0.50	0.94	0.80	1.12	0.93

(Continues)

TABLE 8 (Continued)

Methods	Year	LP	00* U	01 U	02* U	03 U	04 U	05* U	06* U	07* U	08* U	09* U	10 U	00–10 mean
CR-LDSO (Yuan, Cheng, & Yang, 2023)	2023		0.64	0.87	0.81	0.82	0.56	0.65	0.52	0.88	1.14	0.50	0.53	0.72
SDV-LOAM (Yuan, Wang, et al., 2023)			0.50	0.62	0.51	0.56	0.37	0.27	0.28	0.32	0.79	0.47	0.47	0.47
Learning-based methods														
SuMa (scan-scan) (Behley & Stachniss, 2018)	2018		2.11	4.31	2.32	1.63	11.9	1.46	0.89	1.87	2.56	1.99	2.15	2.19
SuMa (scan-model) (Behley & Stachniss, 2018)			0.72	1.77	1.06	0.57	0.39	0.50	0.39	0.37	1.01	0.47	0.69	0.84
SuMa (scan-model) (Behley & Stachniss, 2018)		✓	0.64	1.77	1.23	0.57	0.39	0.42	0.51	0.65	1.15	0.57	0.69	0.89
Velas et al. (Velas et al., 2018)			3.02	4.44	3.42	4.94	1.77	2.35	1.88	1.77	2.89	4.94	3.27	3.12
Lo-Net (Li et al., 2019)	2019		1.47	1.36	1.52	1.03	0.51	1.04	0.71	1.70	2.12	1.37	1.80	1.42
LO-Net + Mapping (Li et al., 2019)			0.78	1.42	1.01	0.73	0.56	0.62	0.55	0.56	1.08	0.77	0.92	0.83
SuMa++ (Chen et al., 2019)		✓	0.64	1.60	1.00	0.67	0.37	0.40	0.46	0.34	1.10	0.47	0.66	0.70
Cho et al. (Cho et al., 2020)	2020		–	–	–	–	–	–	–	–	–	4.87	5.02	4.95
DMLO (Li & Wang, 2020)			0.83	3.14	1.08	0.80	0.71	0.90	0.91	0.58	1.02	1.02	1.08	1.09
DMLO+Mapping (Li & Wang, 2020)			0.73	1.91	1.01	1.06	0.61	0.65	0.65	0.53	0.93	0.58	0.75	0.85
DVL-SLAM (Shin et al., 2020)		✓	0.93	1.47	1.11	0.92	0.67	0.82	0.92	1.26	1.32	0.66	0.70	0.98
LoDoNet (Zheng et al., 2020)			1.43	0.96	1.46	2.12	0.65	1.07	0.62	1.86	2.04	0.63	1.18	1.27

TABLE 8 (Continued)

Methods	Year	LP	00* U	01 U	02* U	03 U	04 U	05* U	06* U	07* U	08* U	09* U	10 U	00–10 mean
DeLORA (Nubert et al., 2021)	2021		–	–	–	–	–	–	–	–	–	6.05	6.44	6.25
PSF-LO (Chen, Wang, et al., 2021)			0.64	1.32	0.87	0.75	0.66	0.45	0.47	0.46	0.94	0.56	0.54	0.74
PSF-LO no-o (Chen, Wang, et al., 2021)			0.65	1.32	0.92	0.76	0.65	0.45	0.48	0.45	0.95	0.55	0.56	0.76
PWCLO-Net (Wang, Wu, Liu, & Wang, 2021)			0.78	0.67	0.86	0.76	0.37	0.45	0.27	0.60	1.26	0.79	1.69	0.77
SA-LOAM-ODOM (Li, Kong, Zhao, Li, et al., 2021)			0.59	1.89	0.77	0.87	0.59	0.45	0.52	0.41	0.85	0.68	0.78	0.76
SA-LOAM-LOOP (Li, Kong, Zhao, Li, et al., 2021)		✓	0.59	1.89	0.79	0.87	0.59	0.37	0.52	0.41	0.84	0.77	0.78	0.76
SelfVoxeLO (Xu et al., 2021)			–	–	–	–	–	–	–	2.51	2.65	2.86	3.22	2.81
Generalized LOAM (Honda et al., 2022)	2022		0.94	–	1.33	–	0.87	–	0.64	–	1.06		0.99	0.97
RSLO (Xu, Lin, et al., 2022)			–	–	–	–	–	–	–	2.37	2.14	2.61	2.33	2.36
TVL-SLAM (Chou & Chou, 2022)		✓	0.88	4.2	1.88	0.74	0.22	0.43	0.30	0.37	2.14	1.14	0.64	1.18
HPPLO-Net (Zhou et al., 2023)	2023		1.01	3.17	1.56	1.04	0.64	0.88	0.56	0.89	1.12	1.45	1.23	1.23
RGB-L (Sauerbeck et al., 2023)			0.70	1.10	0.84	1.03	0.45	0.47	0.88	0.58	1.09	0.87	0.84	0.80
TransLO (Liu et al., 2023)			–	–	–	–	–	–	–	0.55	1.29	0.95	1.18	0.99

Note: U, urban road; H, highway; C, country road; *, scenario with loop closure; –, not available; LP, method with loop closure detection (LCD) module. Red, green and blue fonts denote the first, second and third places, respectively.

TABLE 9 Relative rotational error (%/100 m) on KITTI odometry (Geiger et al., 2012).

Methods	Year	LP	00* U	01 U	02* U	03 U	04 U	05* U	06* U	07* U	08* U	09* U	10 U	00–10 mean
Handcrafted methods														
ICP-po2po (Besl & McKay, 1992)	1992		2.99	2.58	3.39	5.05	4.02	1.93	1.59	3.35	4.93	2.89	4.74	3.40
ICP-po2pl (Besl & McKay, 1992)			1.73	2.58	2.74	1.63	2.58	1.08	1.00	1.42	2.14	1.71	2.60	1.93
GICP (Segal et al., 2009)	2009		0.64	0.91	0.77	1.08	1.07	0.54	0.46	0.45	0.75	0.77	0.62	0.73
CLS (Velas et al., 2016)	2016		0.95	1.05	0.86	1.09	0.71	0.92	0.46	0.73	1.05	0.92	1.28	0.91
LOAM (Ji & Singh, 2017)	2017		0.53	0.55	0.55	0.65	0.50	0.38	0.39	0.50	0.44	0.48	0.57	0.50
LeGO-LOAM (Shan & Englot, 2018)	2018	✓	0.59	0.67	1.57	0.72	0.47	0.42	0.59	0.41	0.52	0.69	0.65	0.66
HDL-Graph-SLAM (Koide et al., 2019)	2019	✓	0.83	0.90	1.78	0.78	0.54	0.87	0.36	1.48	1.01	0.91	1.27	0.98
FALO (García Daza et al., 2020)	2020		0.51	1.35	0.28	0.24	0.33	0.18	–	0.34	–	0.13	0.17	0.39
LiTAMIN (Yokozuka et al., 2020)			0.46	0.45	0.46	0.56	0.47	0.39	0.29	0.34	0.42	0.45	0.90	0.46
LiTAMIN (Yokozuka et al., 2020)		✓	0.41	0.45	0.45	0.56	0.47	0.35	0.32	0.33	0.37	0.43	0.90	0.43
DV-LOAM (Wang, Liu, Wang, et al., 2021)	2021	✓	0.30	0.43	0.33	0.48	0.62	0.30	0.33	0.33	0.32	0.32	0.47	0.38
ELO (Zheng & Zhu, 2021)			0.20	0.13	0.18	0.27	0.15	0.17	0.13	0.16	0.21	0.14	0.19	0.18
F-LOAM (Wang, Wang, Chen, & Xie, 2021)			0.43	0.60	0.52	0.66	0.52	0.36	0.39	0.39	0.46	0.55	0.58	0.49
ISC-LOAM (Wang, Wang, & Xie, 2021)		✓	0.42	0.63	0.54	0.72	0.56	0.37	0.41	0.43	0.50	0.59	0.62	0.52
GLIM (Koide et al., 2021a)		✓	0.18	0.15	0.17	0.33	0.17	0.21	0.10	0.17	0.50	0.17	0.31	0.24
LiTAMIN2 (Yokozuka et al., 2021)			0.42	0.40	0.45	0.43	0.41	0.32	0.23	0.57	0.55	0.32	0.59	0.45
LiTAMIN2 (Yokozuka et al., 2021)		✓	0.28	0.46	0.32	0.48	0.52	0.25	0.34	0.32	0.29	0.40	0.47	0.33
MULLS-LO (Pan et al., 2021)			0.18	0.09	0.17	0.22	0.08	0.17	0.11	0.18	0.25	0.11	0.19	0.16
MULLS-SLAM (Pan et al., 2021)		✓	0.13	0.09	0.13	0.22	0.08	0.07	0.08	0.11	0.17	0.12	0.19	0.13
Vizzo et al. (Vizzo et al., 2021)			0.74	0.68	1.00	0.72	1.10	0.92	0.61	0.42	0.55	0.61	0.66	0.84

TABLE 9 (Continued)

Methods	Year	LP	00* U	01 U	02* U	03 U	04 U	05* U	06* U	07* U	08* U	09* U	10 U	00–10 mean
E-LOAM (Guo et al., 2023)	2022	✓	0.47	0.62	0.82	0.68	0.50	0.36	0.57	0.66	0.63	0.57	0.63	0.59
T-LOAM (Zhou, Guo, et al., 2022)			0.60	0.52	0.39	0.82	0.68	0.32	0.31	0.47	0.33	0.40	0.61	0.49
Learning-based methods														
SuMa (scan-scan) (Behley & Stachniss, 2018)	2018		0.92	1.21	0.78	0.73	1.05	0.76	0.57	1.09	0.95	0.76	0.94	0.88
SuMa (scan-model) (Behley & Stachniss, 2018)			0.30	0.47	0.39	0.46	0.27	0.23	0.14	0.33	0.35	0.25	0.27	0.33
SuMa (scan-model) (Behley & Stachniss, 2018)		✓	0.22	0.47	0.41	0.46	0.27	0.20	0.28	0.52	0.35	0.20	0.27	0.32
Lo-Net (Li et al., 2019)	2019		0.72	0.47	0.71	0.66	0.65	0.69	0.50	0.89	0.77	0.58	0.93	0.71
LO-Net + Mapping (Li et al., 2019)			0.42	0.40	0.45	0.59	0.54	0.35	0.35	0.45	0.43	0.38	0.41	0.43
SuMa++ (Chen et al., 2019)		✓	0.22	0.46	0.37	0.46	0.26	0.20	0.21	0.19	0.35	0.23	0.38	0.29
Cho et al. (Cho et al., 2020)	2020		–	–	–	–	–	–	–	–	–	1.95	1.83	1.89
DMLO (Li & Wang, 2020)			0.44	1.15	0.53	0.49	0.53	0.46	0.46	0.51	0.48	0.45	0.59	0.55
DMLO+Mapping (Li & Wang, 2020)			0.44	0.54	0.58	0.53	0.66	0.33	0.34	0.51	0.48	0.30	0.52	0.47
LoDoNet (Zheng et al., 2020)			0.69	0.28	0.57	0.98	0.45	0.59	0.34	1.64	0.97	0.35	0.45	0.66
DeLORA (Nubert et al., 2021)	2021		–	–	–	–	–	–	–	–	–	62.15	3.00	2.58
PWCLO-Net (Wang, Wu, Liu, & Wang, 2021)			0.42	0.23	0.41	0.44	0.40	0.27	0.22	0.44	0.55	0.35	0.62	0.40
SA-LOAM-ODOM (Li, Kong, Zhao, Li, et al., 2021)			0.25	0.48	0.28	0.46	0.35	0.24	0.25	0.22	0.27	0.28	0.35	0.31
SA-LOAM-LOOP (Li, Kong, Zhao, Li, et al., 2021)		✓	0.18	0.48	0.27	0.46	0.35	0.16	0.24	0.18	0.25	0.25	0.35	0.28
SelfVoxeLO (Xu et al., 2021)			–	–	–	–	–	–	–	1.15	1.00	1.17	1.26	1.15

(Continues)

TABLE 9 (Continued)

Methods	Year	LP	00* U	01 U	02* U	03 U	04 U	05* U	06* U	07* U	08* U	09* U	10 U	00–10 mean
H-VLO (Aydemir et al., 2022)	2022		0.62	0.46	0.60	0.47	0.36	0.35	0.30	0.48	0.38	0.34	0.52	0.44
RSLO (Xu, Lin, et al., 2022)			–	–	–	–	–	–	–	1.15	0.92	1.05	0.94	1.02
TVL-SLAM (Chou & Chou, 2022)		✓	0.45	0.53	0.54	0.71	0.29	0.23	0.19	0.31	0.86	0.53	0.53	0.52
HPPLO-Net (Zhou et al., 2023)	2023		0.55	0.91	0.61	0.82	0.97	0.44	0.31	0.57	0.54	0.61	0.66	0.64
RGB-L (Sauerbeck et al., 2023)			0.26	0.36	0.26	0.26	0.22	0.21	0.47	0.29	0.34	0.87	0.34	0.42
TransLO (Liu et al., 2023)										0.43	0.50	0.46	0.61	0.50

Note: U, urban road; H, highway; C, country road; *, scenario with loop closure; –, not available; LP, method with loop closure detection (LCD) module. Red, green and blue fonts denote the first, second and third places, respectively.

(Zheng & Zhu, 2021) in the third (0.18°).

Commercial SLAM

To help industry units make better equipment selections, Table 10 summarizes representative commercial SLAM solutions and their associated companies, countries, sensor types, application scenarios, and websites.

LiDAR-only solutions

Kaarta develops mobile mapping and localization technology (LOAM) (Ji & Singh, 2017; Zhang & Singh, 2014) based on advanced robotics and delivers it as software as a service (SaaS) and integrated systems. Kitware's versatile library takes a LiDAR frame as input and outputs a pose, which allows researchers to freely replicate/modify into commercial and non-commercial projects under an Apache 2 licence. It is incorporated into a ParaView plugin and offered as a ROS package. SLAMTec's SLAMKit, embedded in a robot's controller through software licensing, enables the robot to map and localize autonomously. Users can effortlessly customize robot applications using standardized software interfaces. GoSLAM commits to delivering a superior user experience in mobile 3D laser measurement systems, which excel in self-localization and incremental 3D mapping across various indoor and outdoor environments. 3irobotix's Robot+Egomobile platform merges AI technology, advanced sensors, and chip processing, substantially boosting the perception and cognition of mobile robots. It delivers outstanding solutions for various robot scenarios. Michael Baker's wearable LiDAR expedites data collection and seamlessly captures indoor and outdoor environments to create comprehensive views of an entire airport.

LI solutions

Google's Cartographer serves as a standalone C++ library in the ROS repository and provides a real-time 2D or 3D solution across multiple platforms and sensor configurations. CRISO's Wildcat SLAM works in conjunction with IMU and LiDAR sensors with the option to add other sensors such as chemical, radiological, gas, GPS, and Wi-Fi. FARO's GeoSLAM develops continuous-time SLAM technology that uses continuous sensor trajectory to generate accurate and detailed 3D maps of a scanned environment. Kolida SLAM-K120 handheld 3D laser scanner integrates advanced LiDAR and IMU in the industry and operates independently of GNSS. It enables a 360° rotation, boasts a 285° FOV, and facilitates effortless acquisition of high-precision point clouds during movement.

LV solutions

GreenValley offers diverse high-precision mapping methods like PPK-SLAM, RTK-SLAM, and pure SLAM, swiftly acquiring point cloud data with absolute coordinates. Paired with the proprietary software LiFuser-BP, LiDAR 360, and LiDAR 360MLS software, it effectively addresses challenges in diverse scenarios such as surveying, mining, forestry, and road component census. FEIMA SLAM100 integrates an industry-level SLAM algorithm with versatile external interfaces and facilitates connections to devices like panoramic cameras, GPS modules, cars, UAVs, and more. It enhances data collection diversity and adaptability across various application scenarios.

TABLE 10 Representative commercial LiDAR SLAM Solutions.

Sensors	Solutions	Company	Country	Applicable scenarios	Website
L	LOAM	Kaarta	USA	Infrastructure, facilities management, geospatial, mining, and forestry industries	https://kaartainc.godaddysites.com/
	Kitware SLAM Library	Kitware	USA	3D modelling, object detection/segmentation, change detection	https://www.kitware.com/lidar-slam-spotlight-on-kitwares-open-source-library/
	SLAMKit	SLAMTEC	China	Cleaning robots, inspection robots, food delivery robots, AGV	https://www.slamtec.com/en
	GoSLAM	GoSLAM	China	Stack measurement, building information collection, digitization of under facilities, agriculture, forestry, and geology	https://www.goslam.com/
	Egomobile	3irobotix	China	2D and 3D environment modelling, 3D semantic maps	http://3irobotics.com/kernel
	Wearable LiDAR	Michael Baker International	USA	All-in-One reality capture, scan to BIM, indoor mapping	https://mbakerintl.com/en/practices/lidar-solutions/
L, I	Cartographer	Google	USA	Service robots, cleaning robots, warehousing robots, autonomous driving	https://github.com/cartographer-project/cartographer
	Wildcat SLAM	CRISO	Australia	Autonomous systems, drones and vehicles, real-time mapping and localization for mobile applications, retro-fitted systems	https://www.csiro.au/en/research/technology-space/robotics/wildcat-slam
	GeoSLAM	FARO	USA	Construction, real estate, mining, engineering, architecture, forestry, security and defence, geospatial, education, media and entertainment	https://geoslam.com/
	SLAM-K120 SLAM-K120 Plus	Kolida	China	Pipeline survey, inertial navigation modelling, underground garage, cultural relics and ancient buildings, mine survey	http://www.kolida.com.cn/productsDetail4-10-265.html

TABLE 10 (Continued)

Sensors	Solutions	Company	Country	Applicable scenarios	Website
L, V	LiFuser-BP	GreenValley	China	Electric power inspection, smart agriculture and forestry, terrain surveying, high-definition map, disaster response, digital city	https://www.lidar360.com/
	SLAM 100-SLAM GO	FEIMA ROBOTICS	China	Topographic mapping, real estate surveying and mapping, digital management, spatial analysis, underground	https://www.feimarobotics.com/en/productDetailSlam100
L, I, V	Grand SLAM	Kudan	Japan	Industrial autonomous mobile robots, consumer robotics, autonomous driving/HD map, drone/UAV surveying and mapping	https://www.kudan.io/our_technology/
	Lixel L1 and L2, LixelStudio	XGRIDS	China	Topographical mapping, agricultural and forestry survey, smart city, engineering survey	https://www.xgrids.cn/lixell1?bd_vid=7765437116044686111
	R100/S100	DASPATIAL	China	High-precision mapping, mine survey, pile survey, garden scanning	https://www.daspatial.com/

Note: I, inertial; L, LiDAR; V, visual.

LIV solutions

Kudan's GrandSLAM accepts a wide range of sensor data such as camera, ToF, LiDAR, IMU, GNSS, and wheel odometry for maximum performance. It integrates all the sensor information into one system rather than just 'filtering' them with accurate time synchronization. XGRIDS introduces Lixel L1 and L2, handheld real-time 3D reconstruction devices, which seamlessly combine visual modules, LiDAR, IMU, and high-performance computer modules. With a user-friendly design, they offer straightforward operation and instant usability. DASPATIAL's R100 and S100, incorporating LiDAR, IMU, and panoramic cameras, operates in handheld and backpack modes for comprehensive data collection. It facilitates continuous collection without pauses, eliminates waiting times at building corners, and supports auxiliary data collection on low-speed vehicles.

FUTURE DIRECTIONS

Multimodal SLAM

Event camera

It is a biomimetic vision sensor that records event streams in time-position-polarity form (Gallego et al., 2020). As a dynamic vision sensor, the event camera has the following advantages compared with traditional cameras: no motion blur, sub-millisecond time delay, and ultra-high dynamic range. These advantages make it have a wide range of application prospects in SLAM. For example, event cameras with extreme motion capture capabilities enable real-time dynamic obstacle avoidance and provide a reliable means for nighttime SLAM. As a new type of sensor, the event camera has attracted extensive attention in the fields of computer vision, robot vision, etc. It has been applied in feature tracking (Kueng et al., 2016), optical flow (Benosman et al., 2013), 3D reconstruction (Matsuda et al., 2015), and SLAM (Censi & Scaramuzza, 2014; Zhou, Gallego, & Shen, 2021; Zihao Zhu et al., 2017). However, due to the uniqueness of event cameras, the processing of noise, spatial–temporal information, and polar data are difficult problems that traditional cameras have never encountered, and all task-level algorithms need to be re-designed (Gallego et al., 2020). Therefore, all existing studies are in the preliminary stage and far from mature, with both challenges and opportunities.

Unconventional sensors

Current SLAM research mainly focuses on sensors such as vision, LiDAR, and IMU. However, signals such as smell, sound, and geomagnetism can also provide important observations for navigation and localization, which have been widely confirmed in nature (Cadena et al., 2016). For example, homing pigeons use their internal geomagnetic sensing cells to determine directions by sensing the Earth's magnetic field. Bees' eyes sense changes in the polarization of sunlight, and their olfactory glands leave odours in their flight paths. Therefore, bees use a combination of vision and smell to achieve precise homing. Bats use echoes to locate and avoid obstacles even in high-speed flights. In addition, Wi-Fi and 5G are also important signal sources for indoor localization. How to build a multi-modal signal source fusion SLAM model? How to judge the confidence of various navigation information sources? And how to automatically identify the noise and outliers of various sources in the back-end optimization? These are the core issues that SLAM urgently needs to break through.

Learning-based SLAM

In 2012, AlexNet (Krizhevsky et al., 2012) won the image recognition championship of ImageNet by a huge margin, and the era of deep learning has come since then. Deep learning has achieved great success in image classification, recognition, segmentation, and other fields, raising the accuracy to an unprecedented level (He, Zhang, et al., 2016; Krizhevsky et al., 2012; Vaswani et al., 2017). Similarly, it has also shown potential in SLAM. For example, new methods have emerged in image matching (DeTone et al., 2018; Xu et al., 2023), point cloud registration (Wang & Solomon, 2019a), semantic segmentation (Milioto et al., 2019), LCD (Cattaneo et al., 2022), and pose estimation (He et al., 2020). In addition, there are even SLAM methods that directly employ end-to-end deep networks (Li et al., 2019; Teed & Deng, 2021). These all prove that deep learning has injected new vitality into the SLAM community, but this does not mean that traditional methods have died. At present, learning-based SLAM methods cannot reach the accuracy and reliability of traditional methods, let alone surpass them. Learning-based SLAM methods have the following problems or trends:

Online learning

This capability is crucial for long-term SLAM systems. Existing deep networks are trained offline on some fixed scene datasets in a closed world. However, SLAM operates in an open environment, and it will constantly encounter new objects, new scenes, and new events, requiring lifelong learning capabilities.

Few-shot learning

Existing deep networks usually rely on a large amount of labelled data and cannot adapt well to challenges from the open world (Cadena et al., 2016). For SLAM, the labelled data required for tasks such as image matching, point cloud registration, LCD, and semantic segmentation are different. Therefore, learning-based SLAM methods are highly dependent on data richness and have a high labelling workload. It is urgent to develop Few-shot learning techniques.

Large models

Nowadays, the research of large models is in full swing, and a number of excellent models have emerged, including GPT (Brown et al., 2020), CLIP (Radford et al., 2021), SAM (Kirillov et al., 2023), etc. Large models have the advantages of strong data processing ability, complex problem-solving ability, high accuracy, and performance. Therefore, the SLAM large model may be one of the ways to solve the above two problems. However, large models require massive data and massive computing resources, which may restrict the development of the SLAM large models.

In short, the current dominance of deep learning in SLAM is weak. How to deeply tap the potential of neural networks in feature expression, function fitting, and scene recognition may be the key to breaking through the bottleneck of SLAM technology.

Multi-robot collaborative SLAM

The scene map is the basis of multi-robot collaborative work, which is a high-precision abstraction and representation of the environment. In a multi-robot system, the scene map acts as a 'smart brain' and provides important

decision support and navigation information for robots. Currently, the technology of 3D mapping based on a single robot is becoming mature. However, limited by the endurance time of a single robot, the efficiency of acquiring 3D information is low and the range is small; limited by the working mode, it is difficult for a single robot to conduct comprehensive real-time analysis of complex structures and scene information; and due to the error accumulation characteristics of SLAM, it is difficult to guarantee the accuracy of long-term and large-scale mapping, which can be solved by dividing a large scene into multiple small scenes.

Unmanned swarms refer to the overall system in which several different forms of intelligent robots cooperate to complete complex tasks within a certain time and space according to the division of tasks (Atanasov et al., 2015; Kegeleirs et al., 2021; Lajoie & Beltrame, 2023; Zhong et al., 2022). It has incomparable advantages over a single robot, but it also faces the following challenges.

Collaborative work

Multi-robots obtain their own swarm poses in a stand-alone and inter-machine collaborative manner and perform path planning and task division so that the robots can share map information and avoid repeated map building.

Data fusion

Unmanned swarms collaborate to explore the environment and build their own local maps. Ultimately, the local maps created by each other need to be globally fused to ensure the consistency and completeness of the environment map.

Dynamic update

Changes may occur at any time in the real environment, such as the movement of obstacles, changes in the environment, etc., requiring timely map updates.

Scholars can start from these challenges to break through the difficulties of cooperative localization, environment sensing, and dynamic updating in multi-robot SLAM, so as to dramatically improve the efficiency, mapping accuracy, stability, and robustness of the single-robot SLAM.

Quantum SLAM

Quantum information science is an emerging frontier discipline based on the theory of quantum mechanics and interdisciplinary integration with communication science, computer science, and other multidisciplinary disciplines (Liu & Hersam, 2019). In 2022, three pioneers of quantum information won the Nobel Prize in Physics, which indicates that quantum information may trigger a new round of technological revolution in the future. Quantum information science has shown great potential in quantum computing, quantum communication, and quantum measurement.

First, quantum computers can break through the computing power limit of classical computers. In the quantum era, SLAM will not be limited by computing resources, and truly realize ultra-large-scale real-time localization and mapping. Quantum communication uses the quantum entanglement effect and superposition state principle to realize information transmission. Its communication speed has an absolute advantage over traditional communication techniques, laying the foundation for multi-robot real-time data sharing. Quantum precision measurement

has a higher anti-interference ability, so it has great advantages in measurement accuracy and stability. Accuracy and stability are the key indicators of SLAM, and SLAM in the quantum era will present unprecedented geometric precision.

In order to actively embrace the future quantum era and grasp the great opportunities, SLAM researchers can plan in advance from the algorithm and hardware level:

- Algorithm: Developing SLAM algorithms based on quantum computing. Quantum computing has attracted the attention of scholars in the field of computer vision, and its applications in tasks such as robust estimation (Chin et al., 2020; Doan et al., 2022) and network structure design (DiAdamo et al., 2021; Zhou, Lv, et al., 2022) have been explored.
- Hardware: Developing quantum SLAM sensors, such as quantum LiDAR and quantum camera, to build a complete SLAM ecosystem with software and hardware.

CONCLUSIONS

This paper offers a comprehensive survey dedicated to LiDAR SLAM, addressing a critical gap in this domain. It extensively discusses the background, encompassing the framework, challenges, and its relation to prior surveys. It provides a comprehensive method taxonomy, briefly outlining their pros and cons. Additionally, it compiles noteworthy open-source LiDAR SLAM datasets, evaluation metrics, experimental performances, and commercial SLAM solutions. Eventually, it highlights promising future directions.

ACKNOWLEDGMENTS

This work was supported in part by National Key R&D Program of China under Grant 2020YFC1521900, in part by National Natural Science Foundation of China (No. 42271444 and 42030102), and in part by the Science and Technology Major Project of Hubei Province under Grant 2021AAA010.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ORCID

Yongjun Zhang https://orcid.org/0000-0001-9845-4251
Pengcheng Shi https://orcid.org/0000-0003-2504-9890
Jiayuan Li https://orcid.org/0000-0002-9850-1668

REFERENCES

Agarwal, S., Mierle, K. & Team, T.C.S. (2022) Ceres solver. Available from: https://github.com/ceres-solver/ceres-solver

Aoki, Y., Goforth, H., Srivatsan, R.A. & Lucey, S. (2019) Pointnetlk: robust & efficient point cloud registration using pointnet. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7156–7165. Available from: https://doi.org/10.1109/CVPR.2019.00733

Atanasov, N., Le Ny, J., Daniilidis, K. & Pappas, G.J. (2015) Decentralized active information acquisition: theory and application to multi-robot SLAM. In: *2015 IEEE international conference on robotics and automation (ICRA)*, pp. 4775–4782. Available from: https://doi.org/10.1109/ICRA.2015.7139863

Aydemir, E., Fetic, N. & Unel, M. (2022) H-VLO: hybrid LiDAR-camera fusion for self-supervised Odometry. In: *2022 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 3302–3307. Available from: https://doi.org/10.1109/IROS47612.2022.9981111

Bai, C., Xiao, T., Chen, Y., Wang, H., Zhang, F. & Gao, X. (2022) Faster-LIO: lightweight tightly coupled LiDAR-inertial odometry using parallel sparse incremental voxels. *IEEE Robotics and Automation Letters*, 7(2), 4861–4868. Available from: https://doi.org/10.1109/LRA.2022.3152830

Bailey, T. & Durrant-Whyte, H. (2006) Simultaneous localization and mapping (SLAM): part II. *IEEE Robotics and Automation Magazine*, 13(3), 108–117. Available from: https://doi.org/10.1109/MRA.2006.1678144

Barath, D., Noskova, J. & Matas, J. (2022) Marginalizing sample consensus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11), 8420–8432. Available from: https://doi.org/10.1109/TPAMI.2021.3103562

Barnes, D., Gadd, M., Murcutt, P., Newman, P. & Posner, I. (2020) The Oxford radar robotcar dataset: a radar extension to the Oxford robotcar dataset. In: *2020 IEEE international conference on robotics and automation (ICRA)*, pp. 6433–6438. Available from: https://doi.org/10.1109/ICRA40945.2020.9196884

Behley, J. & Stachniss, C. (2018) Efficient surfel-based SLAM using 3D laser range data in urban environments. *Robotics: Science and Systems*, 2018, 59. Available from: https://doi.org/10.15607/RSS.2018.XIV.016

Benosman, R., Clercq, C., Lagorce, X., Ieng, S.-H. & Bartolozzi, C. (2013) Event-based visual flow. *IEEE Transactions on Neural Networks and Learning Systems*, 25(2), 407–417. Available from: https://doi.org/10.1109/TNNLS.2013.2273537

Besl, P.J. & McKay, N.D. (1992) A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239–256. Available from: https://doi.org/10.1109/34.121791

Biber, P. & Strasser, W. (2003) The normal distributions transform: a new approach to laser scan matching. *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)*, 3, 2743–2748. Available from: https://doi.org/10.1109/IROS.2003.1249285

Borrmann, D., Elseberg, J., Lingemann, K., Nüchter, A. & Hertzberg, J. (2008) Globally consistent 3D mapping with scan matching. *Robotics and Autonomous Systems*, 56(2), 130–142. Available from: https://doi.org/10.1016/j.robot.2007.07.002

Bosse, M. & Zlot, R. (2009) Continuous 3D scan-matching with a spinning 2D laser. In: *2009 IEEE International Conference on Robotics and Automation*, pp. 4312–4319. Available from https://doi.org/10.1109/ROBOT.2009.5152851

Bouaziz, S., Tagliasacchi, A. & Pauly, M. (2013) Sparse iterative closest point. *Computer Graphics Forum*, 32(5), 113–123. Available from: https://doi.org/10.1111/cgf.12178

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P. et al. (2020) Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901. Available from: https://doi.org/10.5555/3495724.349588

Bry, A., Bachrach, A. & Roy, N. (2012) State estimation for aggressive flight in GPS-denied environments using onboard sensing. In: *2012 IEEE international conference on robotics and automation*, pp. 1–8. Available from: https://doi.org/10.1109/ICRA.2012.6225295

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J. et al. (2016) Past, present, and future of simultaneous localization and mapping: toward the robust-perception age. *IEEE Transactions on Robotics*, 32(6), 1309–1332. Available from: https://doi.org/10.1109/TRO.2016.2624754

Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q. et al. (2020) nuScenes: a multimodal dataset for autonomous driving. In: *2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 11618–11628. Available from: https://doi.org/10.1109/CVPR42600.2020.01164

Cai, H., Ou, N. & Wang, J. (2023) Visual-LiDAR Odometry and mapping with monocular scale correction and motion compensation. *arXiv Prepr. arXiv2304.08978* Available from: https://doi.org/10.48550/arXiv.2304.08978

Campos, C., Elvira, R., Rodríguez, J.J.G., Montiel, J.M.M. & Tardós, J.D. (2021) ORB-SLAM3: an accurate open-source library for visual, visual–inertial, and multimap SLAM. *IEEE Transactions on Robotics*, 37(6), 1874–1890. Available from: https://doi.org/10.1109/TRO.2021.3075644

Cao, K., Liu, R., Wang, Z., Peng, K., Zhang, J., Zheng, J. et al. (2023) Tightly-coupled LiDAR-visual SLAM based on geometric features for mobile agents. In: *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1–8. Available from: https://doi.org/10.1109/ROBIO58561.2023.10354794

Carlevaris-Bianco, N., Ushani, A.K. & Eustice, R.M. (2016) University of Michigan North Campus long-term vision and lidar dataset. *International Journal of Robotics Research*, 35(9), 1023–1035. Available from: https://doi.org/10.1177/0278364915614638

Carlone, L. & Calafiore, G.C. (2018) Convex relaxations for pose graph optimization with outliers. *IEEE Robotics and Automation Letters*, 3(2), 1160–1167. Available from: https://doi.org/10.1109/LRA.2018.2793352

Carlone, L., Calafiore, G.C., Tommolillo, C. & Dellaert, F. (2016) Planar pose graph optimization: duality, optimal solutions, and verification. *IEEE Transactions on Robotics*, 32(3), 545–565. Available from: https://doi.org/10.1109/TRO.2016.2544304

Carlone, L., Censi, A. & Dellaert, F. (2014) Selecting good measurements via ? 1 relaxation: a convex approach for robust estimation over graphs. In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2667–2674. Available from: https://doi.org/10.1109/IROS.2014.6942927

Casafranca, J.J., Paz, L.M. & Piniés, P. (2013) A back-end L1 norm based solution for factor graph SLAM. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 17–23. Available from: https://doi.org/10.1109/IROS.2013.6696326

Cattaneo, D., Vaghi, M. & Valada, A. (2022) Lcdnet: deep loop closure detection and point cloud registration for lidar slam. *IEEE Transactions on Robotics*, 38(4), 2074–2093. Available from: https://doi.org/10.1109/TRO.2022.3150683

Censi, A. (2008) An ICP variant using a point-to-line metric. In: *2008 IEEE international conference on robotics and automation*, pp. 19–25. Available from: https://doi.org/10.1109/ROBOT.2008.4543181

Censi, A. & Scaramuzza, D. (2014) Low-latency event-based visual odometry. In: *2014 IEEE international conference on robotics and automation (ICRA)*, pp. 703–710. Available from: https://doi.org/10.1109/IROS.2016.7758089

Chang, M.F., Lambert, J., Sangkloy, P., Singh, J., Bak, S., Hartnett, A. et al. (2019) Argoverse: 3D tracking and forecasting with rich maps. In: *2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 8740–8749. Available from: https://doi.org/10.1109/CVPR.2019.00895

Chen, G., Wang, B., Wang, X., Deng, H., Wang, B. & Zhang, S. (2021) PSF-LO: parameterized semantic features based Lidar Odometry. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 5056–5062. Available from: https://doi.org/10.1109/ICRA48506.2021.9561554

Chen, K., Nemiroff, R. & Lopez, B.T. (2023) Direct lidar-inertial odometry: lightweight lio with continuous-time motion correction. In: *2023 IEEE international conference on robotics and automation (ICRA)*, pp. 3983–3989. Available from: https://doi.org/10.1109/ICRA48891.2023.10160508

Chen, S., Ma, H., Jiang, C., Zhou, B., Xue, W., Xiao, Z. et al. (2022) NDT-LOAM: a real-time lidar odometry and mapping with weighted NDT and LFA. *IEEE Sensors Journal*, 22(4), 3660–3671. Available from: https://doi.org/10.1109/JSEN.2021.3135055

Chen, S.W., Nardari, G.V., Lee, E.S., Qu, C., Liu, X., Romero, R.A. et al. (2020) SLOAM: semantic lidar odometry and mapping for forest inventory. *IEEE Robotics and Automation Letters*, 5(2), 612–619. Available from: https://doi.org/10.1109/LRA.2019.2963823

Chen, X., Läbe, T., Milioto, A., Röhling, T., Behley, J. & Stachniss, C. (2022) OverlapNet: a siamese network for computing LiDAR scan similarity with applications to loop closing and localization. *Autonomous Robots*, 46(1), 61–81. Available from: https://doi.org/10.1007/s10514-021-09999-0

Chen, X., Milioto, A., Palazzolo, E., Giguère, P., Behley, J. & Stachniss, C. (2019) SuMa++: efficient LiDAR-based semantic SLAM. In: *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 4530–4537. Available from: https://doi.org/10.1109/IROS40897.2019.8967704

Chen, X., Vizzo, I., Läbe, T., Behley, J. & Stachniss, C. (2021) Range image-based LiDAR localization for autonomous vehicles. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 5802–5808. Available from: https://doi.org/10.1109/ICRA48506.2021.9561335

Chen, Y. & Medioni, G. (1992) Object modelling by registration of multiple range images. *Image and Vision Computing*, 10(3), 145–155. Available from: https://doi.org/10.1016/0262-8856(92)90066-C

Chen, Y., Wang, J., Li, J., Lu, C., Luo, Z., Xue, H. et al. (2018) LiDAR-video driving dataset: learning driving policies effectively. In: *2018 IEEE/CVF conference on computer vision and pattern recognition*, pp. 5870–5878. Available from: https://doi.org/10.1109/CVPR.2018.00615

Chetverikov, D., Stepanov, D. & Krsek, P. (2005) Robust Euclidean alignment of 3D point sets: the trimmed iterative closest point algorithm. *Image and Vision Computing*, 23(3), 299–309. Available from: https://doi.org/10.1016/j.imavis.2004.05.007

Chin, T.-J., Suter, D., Ch'ng, S.-F. & Quach, J. (2020) Quantum robust fitting. Available from: https://doi.org/10.1007/978-3-030-69525-5_29

Cho, Y., Kim, G. & Kim, A. (2020) Unsupervised geometry-aware deep LiDAR Odometry. In: *2020 IEEE international conference on robotics and automation (ICRA)*, pp. 2145–2152. Available from: https://doi.org/10.1109/ICRA40945.2020.9197366

Choi, Y., Kim, N., Hwang, S., Park, K., Yoon, J.S., An, K. et al. (2018) KAIST multi-spectral day/night data set for autonomous and assisted driving. *IEEE Transactions on Intelligent Transportation Systems*, 19(3), 934–948. Available from: https://doi.org/10.1109/CVPR42600.2020.01164

Chou, C.-C. & Chou, C.-F. (2022) Efficient and accurate tightly-coupled visual-Lidar SLAM. *IEEE Transactions on Intelligent Transportation Systems*, 23(9), 14509–14523. Available from: https://doi.org/10.1109/TITS.2021.3130089

Choy, C., Dong, W. & Koltun, V. (2020) Deep global registration. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2514–2523. Available from: https://doi.org/10.1109/CVPR42600.2020.00259

Davison, A.J., Reid, I.D., Molton, N.D. & Stasse, O. (2007) MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6), 1052–1067. Available from: https://doi.org/10.1109/TPAMI.2007.1049

Debeunne, C. & Vivet, D. (2020) A review of visual-LiDAR fusion based simultaneous localization and mapping. *Sensors*, 20(7), 2068. Available from: https://doi.org/10.3390/s20072068

Dellaert, F. (2012) Factor graphs and GTSAM: a hands-on introduction. *Georgia Institute of Technology, Tech Report*, 2, 4. Available from: http://hdl.handle.net/1853/45226

Dellenbach, P., Deschaud, J.-E., Jacquet, B. & Goulette, F. (2022) CT-ICP: real-time elastic LiDAR Odometry with loop closure. In: *2022 international conference on robotics and automation (ICRA)*, pp. 5580–5586. Available from: https://doi.org/10.1109/ICRA46639.2022.9811849

Deschaud, J.-E. (2018) IMLS-SLAM: scan-to-model matching based on 3D data. In: *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 2480–2485. Available from: https://doi.org/10.1109/ICRA.2018.8460653

DeTone, D., Malisiewicz, T. & Rabinovich, A. (2018) SuperPoint: self-supervised interest point detection and description. In: *2018 IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW)*, pp. 337–33712. Available from: https://doi.org/10.1109/CVPRW.2018.00060

DiAdamo, S., Nötzel, J., Zanger, B. & Beşe, M.M. (2021) Qunetsim: a software framework for quantum networks. *IEEE Transactions on Quantum Engineering*, 2, 1–12. Available from: https://doi.org/10.1109/TQE.2021.3092395

Doan, A.-D., Sasdelli, M., Suter, D. & Chin, T.-J. (2022) A hybrid quantum-classical algorithm for robust fitting. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 417–427. Available from: https://doi.org/10.1109/CVPR52688.2022.00051

Droeschel, D. & Behnke, S. (2018) Efficient continuous-time SLAM for 3D Lidar-based online mapping. In: *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 5000–5007. Available from: https://doi.org/10.1109/ICRA.2018.8461000

Duan, X., Hu, Q., Zhao, P., Yu, F. & Ai, M. (2023) A low-drift and real-time localisation and mapping method for handheld hemispherical view LiDAR-IMU integration system. *The Photogrammetric Record*, 38(182), 176–196. Available from: https://doi.org/10.1111/phor.12447

Durrant-Whyte, H. & Bailey, T. (2006) Simultaneous localization and mapping: part I. *IEEE Robotics and Automation Magazine*, 13(2), 99–110. Available from: https://doi.org/10.1109/MRA.2006.1638022

Engel, J., Koltun, V. & Cremers, D. (2018) Direct sparse Odometry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(3), 611–625. Available from: https://doi.org/10.1109/TPAMI.2017.2658577

Engel, J., Schoeps, T. & Cremers, D. (2014) LSD-SLAM: large-scale direct monocular SLAM. *Computer Vision – ECCV 2014, PT Ii*, 8690, 834–849. Available from: https://doi.org/10.1007/978-3-319-10605-2_54

Ervan, O. & Temeltas, H. (2023) A histogram-based sampling method for point cloud registration. *The Photogrammetric Record*, 38(183), 210–232. Available from: https://doi.org/10.1111/phor.12448

Fan, T., Wang, H., Rubenstein, M. & Murphey, T. (2020) CPL-SLAM: efficient and certifiably correct planar graph-based SLAM using the complex number representation. *IEEE Transactions on Robotics*, 36(6), 1719–1737. Available from: https://doi.org/10.1109/TRO.2020.3006717

Fischler, M.A. & Bolles, R.C. (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395. Available from: https://doi.org/10.1145/358669.358692

Forster, C., Pizzoli, M. & Scaramuzza, D. (2014) SVO: Fast semi-direct monocular visual odometry. In: *2014 IEEE international conference on robotics and automation (ICRA)*, pp. 15–22. Available from: https://doi.org/10.1109/ICRA.2014.6906584

Fraundorfer, F. & Scaramuzza, D. (2012) Visual odometry part II: matching, robustness, optimization, and applications. *EEE Robotics & Automation Magazine*, 19(2), 78–90. Available from: https://doi.org/10.1109/MRA.2012.2182810

Frosi, M. & Matteucci, M. (2022) ART-SLAM: Accurate Real-Time 6DoF LiDAR SLAM. *IEEE Robotics and Automation Letters*, 7(2), 2692–2699. Available from: https://doi.org/10.1109/LRA.2022.3144795

Gallego, G., Delbrück, T., Orchard, G., Bartolozzi, C., Taba, B., Censi, A. et al. (2020) Eventbased vision: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1), 154–180. Available from: https://doi.org/10.1109/TPAMI.2020.3008413

García Daza, I., Rentero, M., Salinas Maldonado, C., Izquierdo Gonzalo, R., Hernández Parra, N., Ballardini, A. et al. (2020) Fail-aware lidar-based odometry for autonomous vehicles. *Sensors*, 20(15), 4097. Available from: https://doi.org/10.3390/s20154097

Geiger, A., Lenz, P., Stiller, C. & Urtasun, R. (2013) Vision meets robotics: the kitti dataset. *International Journal of Robotics Research*, 32(11), 1231–1237. Available from: https://doi.org/10.48550/arXiv.2004.06320

Geiger, A., Lenz, P. & Urtasun, R. (2012) Are we ready for autonomous driving? The KITTI vision benchmark suite. In: *2012 IEEE conference on computer vision and pattern recognition*, pp. 3354–3361. Available from: https://doi.org/10.1109/CVPR.2012.6248074

Geneva, P., Eckenhoff, K., Yang, Y. & Huang, G. (2018) Lips: Lidar-inertial 3d plane slam. In: *2018 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 123–130. Available from: https://doi.org/10.1109/IROS.2018.8594463

Geyer, J., Kassahun, Y., Mahmudi, M., Ricou, X., Durgesh, R., Chung, A.S. et al. (2020) A2d2: Audi autonomous driving dataset. *arXiv* Prepr., arXiv2004.06320. Available from: https://doi.org/10.1109/TITS.2018.2791533

Giubilato, R., Stürzl, W., Wedler, A. & Triebel, R. (2022) Challenges of SLAM in extremely unstructured environments: the DLR planetary stereo, solid-state LiDAR, inertial dataset. *IEEE Robotics and Automation Letters*, 7(4), 8721–8728. Available from: https://doi.org/10.1109/LRA.2022.3188118

Graeter, J., Wilczynski, A. & Lauer, M. (2018) LIMO: Lidar-monocular visual odometry. In: *2018 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 7872–7879. Available from: https://doi.org/10.1109/IROS.2018.8594394

Grant, W.S., Voorhies, R.C. & Itti, L. (2019) Efficient Velodyne SLAM with point and plane features. *Autonomous Robots*, 43(5), 1207–1224. Available from: https://doi.org/10.1007/s10514-018-9794-6

Grisetti, G., Kümmerle, R., Stachniss, C. & Burgard, W. (2010) A tutorial on graph-based SLAM. *IEEE Intelligent Transportation Systems Magazine*, 2(4), 31–43. Available from: https://doi.org/10.1109/MITS.2010.939925

Grisetti, G., Stachniss, C. & Burgard, W. (2009) Nonlinear constraint network optimization for efficient map learning. *IEEE Transactions on Intelligent Transportation Systems*, 10(3), 428–439. Available from: https://doi.org/10.1109/TITS.2009.2026444

Guadagnino, T., Chen, X., Sodano, M., Behley, J., Grisetti, G. & Stachniss, C. (2022) Fast sparse LiDAR odometry using self-supervised feature selection on intensity images. *IEEE Robotics and Automation Letters*, 7(3), 7597–7604. Available from: https://doi.org/10.1109/LRA.2022.3184454

Guo, H., Zhu, J. & Chen, Y. (2023) E-LOAM: LiDAR Odometry and mapping with expanded local structural information. *IEEE Transactions on Intelligent Vehicles*, 8(2), 1911–1921. Available from: https://doi.org/10.1109/TIV.2022.3151665

He, K., Zhang, X., Ren, S. & Sun, J. (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778. Available from: https://doi.org/10.1109/CVPR.2016.90

He, L., Wang, X. & Zhang, H. (2016) 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. Available from: https://doi.org/10.1109/IROS.2016.7759060

He, Y., Sun, W., Huang, H., Liu, J., Fan, H. & Sun, J. (2020) Pvn3d: a deep point-wise 3d keypoints voting network for 6dof pose estimation. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11632–11641. Available from: https://doi.org/10.1109/CVPR42600.2020.01165

Hess, W., Kohler, D., Rapp, H. & Andor, D. (2016) Real-time loop closure in 2D LIDAR SLAM. In: *2016 IEEE international conference on robotics and automation (ICRA)*, pp. 1271–1278. Available from: https://doi.org/10.1109/ICRA.2016.7487258

Holmes, S.A., Klein, G. & Murray, D.W. (2008) An O (N^2) square root unscented Kalman filter for visual simultaneous localization and mapping. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(7), 1251–1263. Available from: https://doi.org/10.1109/TPAMI.2008.189

Honda, K., Koide, K., Yokozuka, M., Oishi, S. & Banno, A. (2022) Generalized LOAM: LiDAR odometry estimation with trainable local geometric features. *IEEE Robotics and Automation Letters*, 7(4), 12459–12466. Available from: https://doi.org/10.1109/LRA.2022.3219022

Hong, H. & Lee, B.H. (2017) Probabilistic normal distributions transform representation for accurate 3D point cloud registration. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3333–3338. Available from: https://doi.org/10.1109/IROS.2017.8206170

Huang, G.P., Mourikis, A.I. & Roumeliotis, S.I. (2008) Analysis and improvement of the consistency of extended Kalman filter based SLAM. In: *2008 IEEE international conference on robotics and automation*, pp. 473–479. Available from: https://doi.org/10.1109/ROBOT.2008.4543252

Huang, G.P., Mourikis, A.I. & Roumeliotis, S.I. (2013) A quadratic-complexity observability-constrained unscented Kalman filter for SLAM. *IEEE Transactions on Robotics*, 29(5), 1226–1243. Available from: https://doi.org/10.1109/TRO.2013.2267991

Huang, L. (2021) Review on LiDAR-based SLAM techniques. In: *2021 International Conference on signal Processing and Machine Learning (CONF-SPML)*, pp. 163–168. Available from: https://doi.org/10.1109/CONF-SPML54095.2021.00040

Huang, S. & Dissanayake, G. (2007) Convergence and consistency analysis for extended Kalman filter based SLAM. *IEEE Transactions on Robotics*, 23(5), 1036–1049. Available from: https://doi.org/10.1109/TRO.2007.903811

Huang, S., Gojcic, Z., Usvyatsov, M., Wieser, A. & Schindler, K. (2021) PREDATOR: registration of 3D point clouds with low overlap. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4265–4274. Available from: https://doi.org/10.1109/CVPR46437.2021.00425

Huang, S.-S., Ma, Z.-Y., Mu, T.-J., Fu, H. & Hu, S.-M. (2020) Lidar-monocular visual Odometry using point and line features. In: *2020 IEEE international conference on robotics and automation (ICRA)*, pp. 1091–1097. Available from: https://doi.org/10.1109/ICRA40945.2020.9196613

Huang, X., Cheng, X., Geng, Q., Cao, B., Zhou, D., Wang, P. et al. (2018) The ApolloScape dataset for autonomous driving. In: *2018 IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW)*, pp. 1067–10676. Available from: https://doi.org/10.1109/CVPRW.2018.00141

Jeong, J., Cho, Y., Shin, Y.-S., Roh, H. & Kim, A. (2018) Complex urban LiDAR data set. In: *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 6344–6351. Available from: https://doi.org/10.1109/ICRA.2018.8460834

Jeong, J., Cho, Y., Shin, Y.-S., Roh, H. & Kim, A. (2019) Complex urban dataset with multi-level sensors from highly diverse urban environments. *International Journal of Robotics Research*, 38(6), 642–657. Available from: https://doi.org/10.1109/ICRA.2018.8460834

Ji, Z. & Singh, S. (2017) Low-drift and real-time lidar odometry and mapping. *Autonomous Robots*, 41(2), 401–416. Available from: https://doi.org/10.1007/s10514-016-9548-2

Jia, Y., Luo, H., Zhao, F., Jiang, G., Li, Y., Yan, J. et al. (2021) Lvio-fusion: a self-adaptive multi-sensor fusion slam framework using actor-critic method. In: *2021 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 286–293. Available from: https://doi.org/10.1109/IROS51168.2021.9635905

Jiang, B. & Shen, S. (2022) A LiDAR-inertial odometry with principled uncertainty modeling. In: *2022 IEEE/RSJ International conference on intelligent robots and systems (IROS)*, pp. 13292–13299. Available from: https://doi.org/10.1109/IROS47612.2022.9981157

Jiang, H., Dang, Z., Wei, Z., Xie, J., Yang, J. & Salzmann, M. (2023) Robust outlier rejection for 3D registration with Variational Bayes. In: *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1148–1157. Available from: https://doi.org/10.1109/CVPR52729.2023.00117

Jiao, J., Ye, H., Zhu, Y. & Liu, M. (2022) Robust odometry and mapping for multi-LiDAR systems with online extrinsic calibration. *IEEE Transactions on Robotics*, 38(1), 351–371. Available from: https://doi.org/10.1109/TRO.2021.3078287

Jiao, J., Zhu, Y., Ye, H., Huang, H., Yun, P., Jiang, L. et al. (2021) Greedy-based feature selection for efficient LiDAR SLAM. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 5222–5228. Available from: https://doi.org/10.1109/ICRA48506.2021.9561262

Júnior, G.P., Rezende, A.M., Miranda, V.R., Fernandes, R., Azpúrua, H., Neto, A.A. et al. (2022) EKF-LOAM: an adaptive fusion of LiDAR SLAM with wheel odometry and inertial data for confined spaces with few geometric features. *IEEE Transactions on Automation Science and Engineering*, 19(3), 1458–1471. Available from: https://doi.org/10.1109/TASE.2022.3169442

Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J.J. & Dellaert, F. (2012) iSAM2: incremental smoothing and mapping using the Bayes tree. *International Journal of Robotics Research*, 31(2), 216–235. Available from: https://doi.org/10.1177/0278364911430419

Kaess, M., Ranganathan, A. & Dellaert, F. (2008) iSAM: incremental smoothing and mapping. *IEEE Transactions on Robotics*, 24(6), 1365–1378. Available from: https://doi.org/10.1177/0278364911430419

Kalman, R.E. (1960) A new approach to linear filtering and prediction problems. Available from: https://doi.org/10.1115/1.3662552

Kegeleirs, M., Grisetti, G. & Birattari, M. (2021) Swarm slam: challenges and perspectives. *Frontiers in Robotics and AI*, 8, 618268. Available from: https://doi.org/10.3389/frobt.2021.618268

Khattak, S., Nguyen, H., Mascarich, F., Dang, T. & Alexis, K. (2020) Complementary multi–modal sensor fusion for resilient robot pose estimation in subterranean environments. In: *2020 International conference on unmanned aircraft systems (ICUAS)*, pp. 1024–1029. Available from: https://doi.org/10.1109/ICUAS48674.2020.9213865

Kim, B., Jung, C., Shim, D.H. & Agha-mohammadi, A.A. (2023) Adaptive keyframe generation based LiDAR inertial odometry for complex underground environments. In: *2023 IEEE international conference on robotics and automation (ICRA)*, pp. 3332–3338. Available from: https://doi.org/10.1109/ICRA48891.2023.10161207

Kim, G., Choi, S. & Kim, A. (2022) Scan context++: structural place recognition robust to rotation and lateral variations in urban environments. *IEEE Transactions on Robotics*, 38(3), 1856–1874. Available from: https://doi.org/10.1109/TRO.2021.3116424

Kim, G. & Kim, A. (2018) 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. Available from: https://doi.org/10.1109/IROS.2018.8593953

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L. et al. (2023) Segment anything. *arXiv* Prepr. arXiv2304.02643. Available from: https://doi.org/10.48550/arXiv.2304.02643

Klein, G. & Murray, D. (2007) Parallel tracking and mapping for small AR workspaces. In: *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 225–234. Available from: https://doi.org/10.1109/ISMAR.2007.4538852

Koide, K., Miura, J. & Menegatti, E. (2019) A portable three-dimensional LIDAR-based system for long-term and wide-area people behavior measurement. *International Journal of Advanced Robotic Systems*, 16(2), 1729881419841532. Available from: https://doi.org/10.1177/1729881419841532

Koide, K., Yokozuka, M., Oishi, S. & Banno, A. (2021a) Globally consistent 3D LiDAR mapping with GPU-accelerated GICP matching cost factors. *IEEE Robotics and Automation Letters*, 6(4), 8591–8598. Available from: https://doi.org/10.1109/LRA.2021.3113043

Koide, K., Yokozuka, M., Oishi, S. & Banno, A. (2021b) Voxelized GICP for fast and accurate 3D point cloud registration. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 11054–11059. Available from: https://doi.org/10.1109/ICRA48506.2021.9560835

Koide, K., Yokozuka, M., Oishi, S. & Banno, A. (2022) Globally consistent and tightly coupled 3D LiDAR inertial mapping. In: *2022 international conference on robotics and automation (ICRA)*, pp. 5622–5628. Available from: https://doi.org/10.1109/ICRA46639.2022.9812385

Krizhevsky, A., Sutskever, I. & Hinton, G.E. (2012) Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. Available from: https://doi.org/10.1145/3065386

Kueng, B., Mueggler, E., Gallego, G. & Scaramuzza, D. (2016) Low-latency visual odometry using event-based feature tracks. In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 16–23. Available from: https://doi.org/10.1109/IROS.2016.7758089

Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K. & Burgard, W. (2011) G2o: a general framework for graph optimization. In: *2011 IEEE international conference on robotics and automation*, pp. 3607–3613. Available from: https://doi.org/10.1109/ICRA.2011.5979949

Lajoie, P.-Y. & Beltrame, G. (2023) Swarm-slam: sparse decentralized collaborative simultaneous localization and mapping framework for multi-robot systems. *arXiv Prepr. arXiv2301.06230* Available from: https://doi.org/10.48550/arXiv.2301.06230

Lajoie, P.-Y., Hu, S., Beltrame, G. & Carlone, L. (2019) Modeling perceptual aliasing in SLAM via discrete–continuous graphical models. *IEEE Robotics and Automation Letters*, 4(2), 1232–1239. Available from: https://doi.org/10.1109/LRA.2019.2894852

Lee, D., Ryu, S., Yeon, S., Lee, Y., Kim, D., Han, C. et al. (2021) Large-scale localization datasets in crowded indoor spaces. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3226–3235. Available from: https://doi.org/10.1109/CVPR46437.2021.00324

Li, J. (2022) A practical $O(N^2)$ outlier removal method for correspondence-based point cloud registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(8), 3926–3939. Available from: https://doi.org/10.1109/TPAMI.2021.3065021

Li, J., Hu, Q. & Ai, M. (2017) Robust feature matching for geospatial images via an affine-invariant coordinate system. *The Photogrammetric Record*, 32(159), 317–331. Available from: https://doi.org/10.1111/phor.12201

Li, J., Hu, Q. & Ai, M. (2020a) Robust geometric model estimation based on scaled Welsch q-norm. *IEEE Transactions on Geoscience and Remote Sensing*, 58(8), 5908–5921. Available from: https://doi.org/10.1109/TGRS.2020.2972982

Li, J., Hu, Q. & Ai, M. (2020b) RIFT: multi-modal image matching based on radiation-variation insensitive feature transform. *IEEE Transactions on Image Processing*, 29, 3296–3310. Available from: https://doi.org/10.1109/TIP.2019.2959244

Li, J., Hu, Q. & Ai, M. (2020c) GESAC: robust graph enhanced sample consensus for point cloud registration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 363–374. Available from: https://doi.org/10.1016/j.isprsjprs.2020.07.012

Li, J., Hu, Q. & Ai, M. (2021) Point cloud registration based on one-point ransac and scale-annealing biweight estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 59(11), 9716–9729. Available from: https://doi.org/10.1109/TGRS.2020.3045456

Li, J., Hu, Q., Ai, M. & Wang, S. (2021) A geometric estimation technique based on adaptive M-estimators: algorithm and applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 5613–5626. Available from: https://doi.org/10.1109/JSTARS.2021.3078516

Li, J., Hu, Q., Zhang, Y. & Ai, M. (2022) Robust symmetric iterative closest point. *ISPRS Journal of Photogrammetry and Remote Sensing*, 185, 219–231. Available from: https://doi.org/10.1016/j.isprsjprs.2022.01.019

Li, J. & Lee, G.H. (2019) USIP: unsupervised stable interest point detection from 3D point clouds. In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 361–370. Available from: https://doi.org/10.1109/ICCV.2019.00045

Li, J., Shi, P., Hu, Q. & Zhang, Y. (2023a) QGORE: quadratic-time guaranteed outlier removal for point cloud registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9), 11136–11151. Available from: https://doi.org/10.1109/TPAMI.2023.3262780

Li, J., Shi, P., Hu, Q. & Zhang, Y. (2023b) RIFT2: speeding-up RIFT with a new rotation-invariance technique. *arXiv Prepr. arXiv2303.00319* Available from: https://doi.org/10.48550/arXiv.2303.00319

Li, J., Wu, W., Yang, B., Zou, X., Yang, Y., Zhao, X. et al. (2023) WHU-helmet: a helmet-based multisensor SLAM dataset for the evaluation of real-time 3-D mapping in large-scale GNSS-denied environments. *IEEE Transactions on Geoscience and Remote Sensing*, 61, 1–16. Available from: https://doi.org/10.1109/TGRS.2023.3275307

Li, J., Xu, W., Shi, P., Zhang, Y. & Hu, Q. (2022) LNIFT: locally normalized image for rotation invariant multimodal feature matching. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–14. Available from: https://doi.org/10.1109/TGRS.2022.3165940

Li, J., Zhang, Y. & Hu, Q. (2021) Robust estimation in robot vision and photogrammetry: a new model and its applications. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, 137–144. Available from: https://doi.org/10.5194/isprs-annals-V-1-2021-137-2021

Li, J., Zhao, P., Hu, Q. & Ai, M. (2020) Robust point cloud registration based on topological graph and cauchy weighted lq-norm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 160, 244–259. Available from: https://doi.org/10.1016/j.isprsjprs.2019.12.008

Li, J., Zhong, R., Hu, Q. & Ai, M. (2016) Feature-based laser scan matching and its application for indoor mapping. *Sensors*, 16(8), 1265. Available from: https://doi.org/10.3390/s16081265

Li, K., Li, M. & Hanebeck, U.D. (2021) Towards high-performance solid-state-LiDAR-inertial Odometry and mapping. *IEEE Robotics and Automation Letters*, 6(3), 5167–5174. Available from: https://doi.org/10.1109/LRA.2021.3070251

Li, L., Kong, X., Zhao, X., Li, W., Wen, F., Zhang, H. et al. (2021) SA-LOAM: semantic-aided LiDAR SLAM with loop closure. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 7627–7634. Available from: https://doi.org/10.1109/ICRA48506.2021.9560884

Li, Q., Chen, S., Wang, C., Li, X., Wen, C., Cheng, M. et al. (2019) LO-net: deep real-time Lidar odometry. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8465–8474. Available from: https://doi.org/10.1109/CVPR.2019.00867

Li, T., Pei, L., Xiang, Y., Wu, Q., Xia, S., Tao, L. et al. (2020) P 3-LOAM: PPP/LiDAR loosely coupled SLAM with accurate covariance estimation and robust RAIM in urban canyon environment. *IEEE Sensors Journal*, 21(5), 6660–6671. Available from: https://doi.org/10.1109/JSEN.2020.3042968

Li, Z. & Wang, N. (2020) DMLO: deep matching LiDAR odometry. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6010–6017. Available from: https://doi.org/10.1109/IROS45743.2020.9341206

Liao, Y., Xie, J. & Geiger, A. (2023) KITTI-360: a novel dataset and benchmarks for urban scene understanding in 2D and 3D. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(3), 3292–3310. Available from: https://doi.org/10.1177/0278364919843996

Lim, H., Kim, D., Kim, B. & Myung, H. (2023) AdaLIO: robust adaptive LiDAR-inertial Odometry in degenerate indoor environments. *arXiv Prepr. arXiv2304.12577* Available from: https://doi.org/10.48550/arXiv.2304.12577

Lin, J., Liu, X. & Zhang, F. (2020) A decentralized framework for simultaneous calibration, localization and mapping with multiple LiDARs. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4870–4877. Available from: https://doi.org/10.1109/IROS45743.2020.9340790

Lin, J. & Zhang, F. (2020) Loam livox: a fast, robust, high-precision LiDAR odometry and mapping package for LiDARs of small FoV. In: *2020 IEEE international conference on robotics and automation (ICRA)*, pp. 3126–3131. Available from: https://doi.org/10.1109/ICRA40945.2020.9197440

Lin, J. & Zhang, F. (2022a) R3LIVE: a robust, real-time, RGB-colored, LiDAR-inertial-visual tightly-coupled state estimation and mapping package. In: *2022 international conference on robotics and automation (ICRA)*, pp. 10672–10678. Available from: https://doi.org/10.1109/ICRA46639.2022.9811935

Lin, J. & Zhang, F. (2022b) R 3 LIVE++: a robust, real-time, radiance reconstruction package with a tightly-coupled LiDAR-inertial-visual state estimator. *arXiv Prepr. arXiv2209.03666*. Available from: https://doi.org/10.48550/arXiv.2209.03666

Lin, J., Zheng, C., Xu, W. & Zhang, F. (2021) R 2 LIVE: a robust, real-time, LiDAR-inertial-visual tightly-coupled state estimator and mapping. *IEEE Robotics and Automation Letters*, 6(4), 7469–7476. Available from: https://doi.org/10.1109/IROS47612.2022.9981107

Liu, J., Wang, G., Jiang, C., Liu, Z. & Wang, H. (2023) TransLO: a window-based masked point transformer framework for large-scale LiDAR odometry. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(2), 1683–1691. Available from: https://doi.org/10.1609/aaai.v37i2.25256

Liu, X. & Hersam, M.C. (2019) 2D materials for quantum information science. *Nature Reviews Materials*, 4(10), 669–684. Available from: https://doi.org/10.1038/s41578-019-0136-x

Liu, Z. & Zhang, F. (2021) BALM: bundle adjustment for lidar mapping. *IEEE Robotics and Automation Letters*, 6(2), 3184–3191. Available from: https://doi.org/10.1109/LRA.2021.3062815

Lowe, D.G. (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60, 91–110. Available from: https://doi.org/10.1023/B:VISI.0000029664.99615.94

Lu, F., Chen, G., Liu, Y., Qu, Z. & Knoll, A. (2020) Rskdd-net: random sample-based keypoint detector and descriptor. *Advances in Neural Information Processing Systems*, 33, 21297–21308. Available from: https://doi.org/10.48550/arXiv.2010.12394

Lu, W., Wan, G., Zhou, Y., Fu, X., Yuan, P. & Song, S. (2019) DeepVCP: an end-to-end deep neural network for point cloud registration. Available from: https://doi.org/10.1109/ICCV.2019.00010

Lu, W., Zhou, Y., Wan, G., Hou, S. & Song, S. (2019) L3-net: towards learning based LiDAR localization for autonomous driving. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6382–6391. Available from: https://doi.org/10.1109/CVPR.2019.00655

Lv, J., Lang, X., Xu, J., Wang, M., Liu, Y. & Zuo, X. (2023) Continuous-time fixed-lag smoothing for LiDAR-inertial-camera SLAM. *IEEE/ASME Transactions on Mechatronics*, 28(4), 2259–2270. Available from: https://doi.org/10.1109/TMECH.2023.3241398

Ma, J., Zhang, J., Xu, J., Ai, R., Gu, W. & Chen, X. (2022) OverlapTransformer: an efficient and yaw-angle-invariant transformer network for LiDAR-based place recognition. *IEEE Robotics and Automation Letters*, 7(3), 6958–6965. Available from: https://doi.org/10.1109/LRA.2022.3178797

Maier-Hein, L., Franz, A.M., dos Santo, T.R., Schmidt, M., Fangerau, M., Meinzer, H.P. et al. (2011) Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(8), 1520–1532. Available from: https://doi.org/10.1109/TPAMI.2011.248

Matsuda, N., Cossairt, O. & Gupta, M. (2015) Mc3d: Motion contrast 3d scanning. In: *2015 IEEE international conference on computational photography (ICCP)*, pp. 1–10. Available from: https://doi.org/10.1109/ICCPHOT.2015.7168370

Milioto, A., Vizzo, I., Behley, J. & Stachniss, C. (2019) RangeNet ++: fast and accurate LiDAR semantic segmentation. In: *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 4213–4220. Available from: https://doi.org/10.1109/IROS40897.2019.8967762

Millane, A., Taylor, Z., Oleynikova, H., Nieto, J., Siegwart, R. & Cadena, C. (2018) C-blox: a scalable and consistent TSDF-based dense mapping approach. In: *2018 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 995–1002. Available from: https://doi.org/10.1109/IROS.2018.8593427

Moosmann, F. & Stiller, C. (2011) Velodyne SLAM. In: *2011 IEEE intelligent vehicles symposium (IV)*, pp. 393–398. Available from: https://doi.org/10.1109/IVS.2011.5940396

Mourikis, A.I. & Roumeliotis, S.I. (2007) A multi-state constraint Kalman filter for vision-aided inertial navigation. In: *Proceedings 2007 IEEE international conference on robotics and automation*, pp. 3565–3572. Available from: https://doi.org/10.1109/ROBOT.2007.364024

Mur-Artal, R., Montiel, J.M.M. & Tardós, J.D. (2015) ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5), 1147–1163. Available from: https://doi.org/10.1109/TRO.2015.2463671

Mur-Artal, R. & Tardós, J.D. (2017) ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Transactions on Robotics*, 33(5), 1255–1262. Available from: https://doi.org/10.1109/TRO.2017.2705103

Nasiri, S.-M., Hosseini, R. & Moradi, H. (2021) Novel parameterization for gauss–Newton methods in 3-D pose graph optimization. *IEEE Transactions on Robotics*, 37(3), 780–797. Available from: https://doi.org/10.1109/TRO.2020.3034021

Nguyen, T.-M., Yuan, S., Cao, M., Lyu, Y., Nguyen, T.H. & Xie, L. (2022) Ntu viral: a visual-inertial-ranging-lidar dataset, from an aerial vehicle viewpoint. *International Journal of Robotics Research*, 41(3), 270–280. Available from: https://doi.org/10.1109/CVPR.2019.00895

Nicolai, A., Skeele, R., Eriksen, C. & Hollinger, G.A. (2016a) Deep learning for laser based Odometry estimation. Available from: https://api.semanticscholar.org/CorpusID:30369588

Nicolai, A., Skeele, R., Eriksen, C. & Hollinger, G.A. (2016b) Deep learning for laser based odometry estimation. In: *RSS Workshop limits and potentials of deep learning in robotics*. Available from: https://api.semanticscholar.org/CorpusID:1658079

Nister, D., Naroditsky, O. & Bergen, J. (2004) Visual odometry. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004*, 1, 1. Available from: https://doi.org/10.1109/CVPR.2004.1315094

Nubert, J., Khattak, S. & Hutter, M. (2021) Self-supervised learning of LiDAR Odometry for robotic applications. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 9601–9607. Available from: https://doi.org/10.1109/ICRA48506.2021.9561063

Olson, E. & Agarwal, P. (2013) Inference on networks of mixtures for robust robot mapping. *International Journal of Robotics Research*, 32(7), 826–840. Available from: https://doi.org/10.1177/0278364913479413

Olson, E., Leonard, J. & Teller, S. (2006) Fast iterative alignment of pose graphs with poor initial estimates. In: *Proceedings 2006 IEEE international conference on robotics and automation, 2006. ICRA 2006*, pp. 2262–2269. Available from: https://doi.org/10.1109/ROBOT.2006.1642040

Pan, Y., Xiao, P., He, Y., Shao, Z. & Li, Z. (2021) MULLS: Versatile LiDAR SLAM via Multi-metric Linear Least Square. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 11633–11640. Available from: https://doi.org/10.1109/ICRA48506.2021.9561364

Pandey, G., McBride, J.R. & Eustice, R.M. (2011) Ford campus vision and lidar data set. *International Journal of Robotics Research*, 30(13), 1543–1552. Available from: https://doi.org/10.1177/0278364911400640

Park, C., Moghadam, P., Kim, S., Elfes, A., Fookes, C. & Sridharan, S. (2018) Elastic LiDAR fusion: dense map-centric continuous-time SLAM. In: *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 1206–1213. Available from: https://doi.org/10.1109/ICRA.2018.8462915

Parra Bustos, A. & Chin, T.-J. (2018) Guaranteed outlier removal for point cloud registration with correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12), 2868–2882. Available from: https://doi.org/10.1109/TPAMI.2023.3262780

Pavlov, A.L., Ovchinnikov, G.W.V., Derbyshev, D.Y., Tsetserukou, D. & Oseledets, I.V. (2018) AA-ICP: iterative closest point with Anderson acceleration. In: *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 3407–3412. Available from: https://doi.org/10.1109/ICRA.2018.8461063

Pitropov, M., Garcia, D.E., Rebello, J., Smart, M., Wang, C., Czarnecki, K. et al. (2021) Canadian adverse driving conditions dataset. *International Journal of Robotics Research*, 40(4–5), 681–690. Available from: https://doi.org/10.1109/CVPR.2018.00803

Qin, C., Ye, H., Pranata, C.E., Han, J., Zhang, S. & Liu, M. (2020) Lins: a lidar-inertial state estimator for robust and efficient navigation. In: *2020 IEEE international conference on robotics and automation (ICRA)*, pp. 8899–8906. Available from: https://doi.org/10.1109/ICRA40945.2020.9197567

Qin, T., Li, P. & Shen, S. (2018) Vins-mono: a robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34(4), 1004–1020. Available from: https://doi.org/10.1109/TRO.2018.2853729

Quenzel, J. & Behnke, S. (2021) Real-time multi-adaptive-resolution-Surfel 6D LiDAR odometry using continuous-time trajectory optimization. In: *2021 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 5499–5506. Available from: https://doi.org/10.1109/IROS51168.2021.9636763

Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S. et al. (2021) Learning transferable visual models from natural language supervision. In: *International conference on machine learning*, pp. 8748–8763. Available from: https://doi.org/10.48550/arXiv.2103.00020

Ramanishka, V., Chen, Y.-T., Misu, T. & Saenko, K. (2018) Toward driving scene understanding: a dataset for learning driver behavior and causal reasoning. In: *2018 IEEE/CVF conference on computer vision and pattern recognition*, pp. 7699–7707. Available from: https://doi.org/10.1109/CVPR.2018.00803

Ramezani, M., Wang, Y., Camurri, M., Wisth, D., Mattamala, M. & Fallon, M. (2020) The newer college dataset: handheld LiDAR, inertial and vision with ground truth. In: *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 4353–4360. Available from: https://doi.org/10.1109/IROS45743.2020.9340849

Reinke, A., Palieri, M., Morrell, B., Chang, Y., Ebadi, K., Carlone, L. et al. (2022) LOCUS 2.0: robust and computationally efficient lidar odometry for real-time 3D mapping. *IEEE Robotics and Automation Letters*, 7(4), 9043–9050. Available from: https://doi.org/10.1109/LRA.2022.3181357

Ros, G., Sellart, L., Materzynska, J., Vazquez, D. & Lopez, A.M. (2016) The SYNTHIA dataset: a large collection of synthetic images for semantic segmentation of urban scenes. In: *2016 IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 3234–3243. Available from: https://doi.org/10.1109/CVPR.2016.352

Rosen, D.M., DuHadway, C. & Leonard, J.J. (2015) A convex relaxation for approximate global optimization in simultaneous localization and mapping. In: *2015 IEEE international conference on robotics and automation (ICRA)*, pp. 5822–5829. Available from: https://doi.org/10.1109/ICRA.2015.7140014

Rozenberszki, D. & Majdik, A.L. (2020) LOL: Lidar-only Odometry and localization in 3D point cloud maps. In: *2020 IEEE international conference on robotics and automation (ICRA)*, pp. 4379–4385. Available from: https://doi.org/10.1109/ICRA40945.2020.9197450

Rusinkiewicz, S. (2019) A symmetric objective function for ICP. *ACM Transactions on Graphics*, 38(4), 1–7. Available from: https://doi.org/10.1145/3306346.3323037

Rusinkiewicz, S. & Levoy, M. (2001) Efficient variants of the ICP algorithm. In: *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pp. 145–152. Available from: https://doi.org/10.1109/IM.2001.924423

Rusu, R.B. & Cousins, S. (2011) 3D is here: point cloud library (PCL). In: *2011 IEEE international conference on robotics and automation*, pp. 1–4. Available from: https://doi.org/10.1109/ICRA.2011.5980567

Saputra, M.R.U., Markham, A. & Trigoni, N. (2018) Visual SLAM and structure from motion in dynamic environments: a survey. *ACM Computing Surveys*, 51(2), 1–36. Available from: https://doi.org/10.1145/3177853

Sauerbeck, F., Obermeier, B., Rudolph, M. & Betz, J. (2023) RGB-L: enhancing indirect visual SLAM using LiDAR-based dense depth maps. In: *2023 3rd international conference on computer, control and robotics (ICCCR)*, pp. 95–100. Available from: https://doi.org/10.1109/ICCCR56747.2023.10194045

Scaramuzza, D. & Fraundorfer, F. (2011) Visual odometry [tutorial]. *IEEE Robotics and Automation Magazine*, 18(4), 80–92. Available from: https://doi.org/10.1109/MRA.2011.943233

Schaupp, L., Bürki, M., Dubé, R., Siegwart, R. & Cadena, C. (2019) OREOS: oriented recognition of 3D point clouds in outdoor scenarios. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3255–3261. Available from: https://doi.org/10.1109/IROS40897.2019.8968094

Schmiedel, T., Einhorn, E. & Gross, H.-M. (2015) IRON: a fast interest point descriptor for robust NDT-map matching and its application to robot localization. In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3144–3151. Available from: https://doi.org/10.1109/IROS.2015.7353812

Schnabel, R. & Klein, R. (2006) Octree-based point-cloud compression. *PBG@ SIGGRAPH*, 3, 111–121. Available from: https://doi.org/10.5555/2386388.2386404

Segal, A., Haehnel, D. & Thrun, S. (2009) Generalized-icp. *Robotics: Science and Systems*, 2(4), 435. Available from: https://doi.org/10.15607/rss.2009.v.021

Seo, Y. & Chou, C.-C. (2019) A tight coupling of vision-Lidar measurements for an effective Odometry. In: *2019 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1118–1123. Available from: https://doi.org/10.1109/IVS.2019.8814164

Shan, T. & Englot, B. (2018) 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)*, pp. 4758–4765. Available from: https://doi.org/10.1109/IROS.2018.8594299

Shan, T., Englot, B., Meyers, D., Wang, W., Ratti, C. & Rus, D. (2020) LIO-SAM: tightly-coupled Lidar inertial Odometry via smoothing and mapping. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5135–5142. Available from: https://doi.org/10.1109/IROS45743.2020.9341176

Shan, T., Englot, B., Ratti, C. & Rus, D. (2021) Lvi-sam: tightly-coupled lidar-visual-inertial odometry via smoothing and mapping. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 5692–5698. Available from: https://doi.org/10.1109/ICRA48506.2021.9561996

Shao, W., Vijayarangan, S., Li, C. & Kantor, G. (2019) Stereo visual inertial lidar simultaneous localization and mapping. In: *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 370–377. Available from: https://doi.org/10.1109/IROS40897.2019.8968012

Shi, P., Li, J. & Zhang, Y. (2023a) LiDAR localization at 100 FPS: a map-aided and template descriptor-based global method. *International Journal of Applied Earth Observation and Geoinformation*, 120, 103336. Available from: https://doi.org/10.1016/j.jag.2023.103336

Shi, P., Li, J. & Zhang, Y. (2023b) A fast LiDAR place recognition and localization method by fusing local and global search. *ISPRS Journal of Photogrammetry and Remote Sensing*, 202, 637–651. Available from: https://doi.org/10.1016/j.isprsjprs.2023.07.008

Shi, P., Ye, Q., Shaoming, Z. & Haifeng, D. (2021) Localization initialization for multi-beam LiDAR considering indoor scene feature. *Acta Geodaetica et Cartographica Sinica*, 50, 1594–1604. Available from: https://doi.org/10.11947/j.AGCS.2021.20210268

Shi, P., Ye, Q. & Zeng, L. (2020) A novel indoor structure extraction based on dense point cloud. *ISPRS International Journal of Geo-Information*, 9(11), 660. Available from: https://doi.org/10.3390/ijgi9110660

Shi, P., Zhang, Y. & Li, J. (2023) LiDAR-based place recognition for autonomous driving: a survey. *arXiv Prepr. arXiv2306.10561* Available from: https://doi.org/10.48550/arXiv.2306.10561

Shin, Y.-S., Park, Y.S. & Kim, A. (2020) DVL-SLAM: sparse depth enhanced direct visual-LiDAR SLAM. *Autonomous Robots*, 44(2), 115–130. Available from: https://doi.org/10.1007/s10514-019-09881-0

Shu, C. & Luo, Y. (2022) Multi-modal feature constraint based tightly coupled monocular visual-lidar odometry and mapping. *IEEE Transactions on Intelligent Vehicles*, 8(5), 3384–3393. Available from: https://doi.org/10.1109/TIV.2022.3215141

Song, Z., Lu, J., Yao, Y. & Zhang, J. (2021) Self-supervised depth completion from direct visual-LiDAR odometry in autonomous driving. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 11654–11665. Available from: https://doi.org/10.1109/TITS.2021.3106055

Sturm, J., Engelhard, N., Endres, F., Burgard, W. & Cremers, D. (2012) A benchmark for the evaluation of RGB-D SLAM systems. In: *2012 IEEE/RSJ international conference on intelligent robots and systems*, pp. 573–580. Available from: https://doi.org/10.1109/IROS.2012.6385773

Su, Y., Wang, T., Shao, S., Yao, C. & Wang, Z. (2021) GR-LOAM: LiDAR-based sensor fusion SLAM for ground robots on complex terrain. *Robotics and Autonomous Systems*, 140, 103759. Available from: https://doi.org/10.1016/j.robot.2021.103759

Tagliabue, A., Tordesillas, J., Cai, X., Santamaria-Navarro, A., How, J.P., Carlone, L. et al. (2021) Lion: Lidar-inertial observability-aware navigator for vision-denied environments. In: *Experimental robotics: The 17th international symposium*, pp. 380–390. Available from: https://doi.org/10.1007/978-3-030-71151-1_34

Taketomi, T., Uchiyama, H. & Ikeda, S. (2017) Visual SLAM algorithms: a survey from 2010 to 2016. *IPSJ Transactions on Computer Vision and Applications*, 9(1), 1–11. Available from: https://doi.org/10.1186/s41074-017-0027-2

Teed, Z. & Deng, J. (2021) Droid-slam: deep visual slam for monocular, stereo, and rgb-d cameras. *Advances in Neural Information Processing Systems*, 34, 16558–16569. Available from: https://doi.org/10.48550/arXiv.2108.10869

Tong, C.H. & Barfoot, T.D. (2013) Gaussian process gauss-Newton for 3D laser-based visual Odometry. In: *2013 IEEE international conference on robotics and automation*, pp. 5204–5211. Available from: https://doi.org/10.1109/ICRA.2013.6631321

Tong, C.H., Furgale, P. & Barfoot, T.D. (2012) Gaussian process gauss-Newton: non-parametric state estimation. In: *2012 Ninth conference on computer and robot vision*, pp. 206–213. Available from: https://doi.org/10.1109/CRV.2012.35

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N. et al. (2017) Attention is all you need. In: *2017 Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPs)*, pp. 6000–6010. Available from: https://doi.org/10.5555/3295222.3295349

Velas, M., Spanel, M. & Herout, A. (2016) Collar line segments for fast odometry estimation from Velodyne point clouds. In: *2016 IEEE international conference on robotics and automation (ICRA)*, pp. 4486–4495. Available from: https://doi.org/10.1109/ICRA.2016.7487648

Velas, M., Spanel, M., Hradis, M. & Herout, A. (2018) CNN for IMU assisted odometry estimation using velodyne LiDAR. In: *2018 IEEE international conference on autonomous robot systems and competitions (ICARSC)*, pp. 71–77. Available from: https://doi.org/10.1109/ICARSC.2018.8374163

Vizzo, I., Chen, X., Chebrolu, N., Behley, J. & Stachniss, C. (2021) Poisson surface reconstruction for LiDAR odometry and mapping. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 5624–5630. Available from: https://doi.org/10.1109/ICRA48506.2021.9562069

Vizzo, I., Guadagnino, T., Mersch, B., Wiesmann, L., Behley, J. & Stachniss, C. (2023) KISS-ICP: in defense of point-to-point ICP – simple, accurate, and robust registration if done the right way. *IEEE Robotics and Automation Letters*, 8(2), 1029–1036. Available from: https://doi.org/10.1109/LRA.2023.3236571

Wan, E.A. & Van der Merwe, R. (2000) The unscented Kalman filter for nonlinear estimation. In: *Proceedings of the IEEE 2000 adaptive systems for signal processing, communications, and control symposium (Cat. No.00EX373)*, pp. 153–158. Available from: https://doi.org/10.1109/ASSPCC.2000.882463

Wang, C., Cao, Z., Li, J., Yu, J. & Wang, S. (2023) Hierarchical distribution-based tightly-coupled LiDAR inertial odometry. *IEEE Transactions on Intelligent Vehicles*, 9(1), 1423–1435. Available from: https://doi.org/10.1109/TIV.2023.3273288

Wang, G., Wu, X., Liu, Z. & Wang, H. (2021) PWCLO-net: deep LiDAR Odometry in 3D point clouds using hierarchical embedding mask optimization. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15905–15914. Available from: https://doi.org/10.1109/CVPR46437.2021.01565

Wang, H., Wang, C., Chen, C.-L. & Xie, L. (2021) F-LOAM: fast LiDAR Odometry and mapping. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4390–4396. Available from: https://doi.org/10.1109/IROS51168.2021.9636655

Wang, H., Wang, C. & Xie, L. (2021) Intensity-SLAM: intensity assisted localization and mapping for large scale environment. *IEEE Robotics and Automation Letters*, 6(2), 1715–1721. Available from: https://doi.org/10.1109/LRA.2021.3059567

Wang, J. & Olson, E. (2014) Robust pose graph optimization using stochastic gradient descent. In: *2014 IEEE international conference on robotics and automation (ICRA)*, pp. 4284–4289. Available from: https://doi.org/10.1109/ICRA.2014.6907482

Wang, W., Liu, J., Wang, C., Luo, B. & Zhang, C. (2021) DV-LOAM: direct visual lidar odometry and mapping. *Remote Sensing*, 13(16), 3340. Available from: https://doi.org/10.3390/rs13163340

Wang, W., Saputra, M.R., Zhao, P., Gusmao, P., Yang, B., Chen, C. et al. (2019) DeepPCO: end-to-end point cloud Odometry through deep parallel neural network. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3248–3254. Available from: https://doi.org/10.1109/IROS40897.2019.8967756

Wang, Y. & Solomon, J. (2019a) Deep closest point: learning representations for point cloud registration. In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 3522–3531. Available from: https://doi.org/10.1109/ICCV.2019.00362

Wang, Y. & Solomon, J. (2019b) PRNet: self-supervised learning for partial-to-partial registration. In: *2019 International Conference on Neural Information Processing Systems (NIPS)*, pp. 8814-8826. Available from: https://doi.org/10.5555/3454287.3455078

Wang, Y., Song, W., Zhang, Y., Huang, F., Tu, Z. & Lou, Y. (2021) MetroLoc: metro vehicle mapping and localization with LiDAR-camera-inertial integration. *arXiv Prepr. arXiv2111.00762* Available from: https://doi.org/10.48550/arXiv.2111.00762

Wang, Y., Sun, Z., Xu, C.-Z., Sarma, S.E., Yang, J. & Kong, H. (2020) LiDAR Iris for loop-closure detection. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5769–5775. Available from: https://doi.org/10.1109/IROS45743.2020.9341010

Wang, Z., Zhang, J., Chen, S., Yuan, C., Zhang, J. & Zhang, J. (2019) Robust high accuracy visual-inertial-laser slam system. In: *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 6636–6641. Available from: https://doi.org/10.1109/IROS40897.2019.8967702

Wang, Z., Zhang, L., Shen, Y. & Zhou, Y. (2022) D-liom: tightly-coupled direct lidar-inertial odometry and mapping. *IEEE Transactions on Multimedia*, 25, 3905–3920. Available from: https://doi.org/10.1109/TMM.2022.3168423

Wiesmann, L., Marcuzzi, R., Stachniss, C. & Behley, J. (2022) Retriever: point cloud retrieval in compressed 3D maps. In: *2022 international conference on robotics and automation (ICRA)*, pp. 10925–10932. Available from: https://doi.org/10.1109/ICRA46639.2022.9811785

Wisth, D., Camurri, M., Das, S. & Fallon, M. (2021) Unified multi-modal landmark tracking for tightly coupled lidar-visual-inertial odometry. *IEEE Robotics and Automation Letters*, 6(2), 1004–1011. Available from: https://doi.org/10.1109/LRA.2021.3056380

Wisth, D., Camurri, M. & Fallon, M. (2022) VILENS: visual, inertial, lidar, and leg odometry for all-terrain legged robots. *IEEE Transactions on Robotics*, 39(1), 309–326. Available from: https://doi.org/10.1109/TRO.2022.3193788

Xiang, Z., Yu, J., Li, J. & Su, J. (2019) Vilivo: virtual lidar-visual odometry for an autonomous vehicle with a multi-camera system. In: *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 2486–2492. Available from: https://doi.org/10.1109/IROS40897.2019.8968484

Xiao, H., Han, Y., Zhao, J., Cui, J., Xiong, L. & Yu, Z. (2022) LIO-vehicle: a tightly-coupled vehicle dynamics extension of LiDAR inertial odometry. *IEEE Robotics and Automation Letters*, 7(1), 446–453. Available from: https://doi.org/10.1109/LRA.2021.3126336

Xu, W., Cai, Y., He, D., Lin, J. & Zhang, F. (2022) FAST-LIO2: Fast direct LiDAR-inertial Odometry. *IEEE Transactions on Robotics*, 38(4), 2053–2073. Available from: https://doi.org/10.1109/TRO.2022.3141876

Xu, W., Yuan, X., Hu, Q. & Li, J. (2023) SAR-optical feature matching: a large-scale patch dataset and a deep local descriptor. *International Journal of Applied Earth Observation and Geoinformation*, 122, 103433. Available from: https://doi.org/10.1016/j.jag.2023.103433

Xu, W. & Zhang, F. (2021) FAST-LIO: a fast, robust LiDAR-inertial Odometry package by tightly-coupled iterated Kalman filter. *IEEE Robotics and Automation Letters*, 6(2), 3317–3324. Available from: https://doi.org/10.1109/LRA.2021.3064227

Xu, Y., Huang, Z., Lin, K.Y., Zhu, X., Shi, J., Bao, H. et al. (2021) Selfvoxelo: sSelf-supervised lidar odometry with voxel-based deep neural networks. In: *Conference on robot learning*, pp. 115–125. Available from: https://doi.org/10.48550/arXiv.2010.09343

Xu, Y., Lin, J., Shi, J., Zhang, G., Wang, X. & Li, H. (2022) Robust self-supervised LiDAR Odometry via representative structure discovery and 3D inherent error modeling. *IEEE Robotics and Automation Letters*, 7(2), 1651–1658. Available from: https://doi.org/10.1109/LRA.2022.3140794

Yang, H., Shi, J. & Carlone, L. (2021) TEASER: fast and certifiable point cloud registration. *IEEE Transactions on Robotics*, 37(2), 314–333. Available from: https://doi.org/10.1109/TRO.2020.3033695

Ye, H., Chen, Y. & Liu, M. (2019) Tightly coupled 3d lidar inertial odometry and mapping. In: *2019 international conference on robotics and automation (ICRA)*, pp. 3144–3150. Available from: https://doi.org/10.1109/ICRA.2019.8793511

Ye, Q., Shi, P., Xu, K., Gui, P. & Zhang, S. (2020) A novel loop closure detection approach using simplified structure for low-cost LiDAR. *Sensors*, 20(8), 2299. Available from: https://doi.org/10.3390/s20082299

Yin, H., Tang, L., Ding, X., Wang, Y. & Xiong, R. (2020) Cae-lo: Lidar odometry leveraging fully unsupervised convolutional auto-encoder for interest point detection and feature description. *arXiv Prepr. arXiv2001.01354*. Available from: https://doi.org/10.48550/arXiv.2001.01354

Yin, H., Tang, L., Ding, X., Wang, Y. & Xiong, R. (2018) LocNet: global localization in 3D point clouds for Mobile vehicles. In: *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 728–733. Available from: https://doi.org/10.1109/IVS.2018.8500682

Yin, P., Wang, F., Egorov, A., Hou, J., Jia, Z. & Han, J. (2022) Fast sequence-matching enhanced viewpoint-invariant 3-D place recognition. *IEEE Transactions on Industrial Electronics*, 69(2), 2127–2135. Available from: https://doi.org/10.1109/TIE.2021.3057025

Yokozuka, M., Koide, K., Oishi, S. & Banno, A. (2020) LiTAMIN: LiDAR-based tracking and mapping by stabilized ICP for geometry approximation with Normal distributions. In: *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 5143–5150. Available from: https://doi.org/10.1109/IROS45743.2020.9341341

Yokozuka, M., Koide, K., Oishi, S. & Banno, A. (2021) LiTAMIN2: ultra light LiDAR-based SLAM using geometric approximation applied with KL-divergence. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 11619–11625. Available from: https://doi.org/10.1109/ICRA48506.2021.9560947

You, Y., Lou, Y., Li, C., Cheng, Z., Li, L., Ma, L. et al. (2020) KeypointNet: a large-scale 3D Keypoint dataset aggregated from numerous human annotations. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13644–13653. Available from: https://doi.org/10.1109/CVPR42600.2020.01366

Yousif, K., Bab-Hadiashar, A. & Hoseinnezhad, R. (2015) An overview to visual odometry and visual SLAM: applications to mobile robotics. *Intelligent Industrial Systems*, 1(4), 289–311. Available from: https://doi.org/10.1007/s40903-015-0032-7

Yuan, Z., Cheng, J. & Yang, X. (2023) CR-LDSO: direct sparse LiDAR-assisted visual odometry with cloud reusing. *IEEE Transactions on Multimedia*, 25, 9397–9409. Available from: https://doi.org/10.1109/TMM.2023.3252161

Yuan, Z., Wang, Q., Cheng, K., Hao, T. & Yang, X. (2023) SDV-LOAM: semi-direct visual–LiDAR Odometry and mapping. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9), 11203–11220. Available from: https://doi.org/10.1109/TPAMI.2023.3262817

Zhang, J., Kaess, M. & Singh, S. (2017) A real-time method for depth enhanced visual odometry. *Autonomous Robots*, 41, 31–43. Available from: https://doi.org/10.1007/s10514-015-9525-1

Zhang, J. & Singh, S. (2014) LOAM: Lidar odometry and mapping in real-time. *Robotics: Science and Systems*, 2(9), 1–9. Available from: https://doi.org/10.15607/RSS.2014.X.007

Zhang, J. & Singh, S. (2015) Visual-lidar odometry and mapping: low-drift, robust, and fast. In: *2015 IEEE international conference on robotics and automation (ICRA)*, pp. 2174–2181. Available from: https://doi.org/10.1109/ICRA.2015.7139486

Zhang, J. & Singh, S. (2018) Laser–visual–inertial odometry and mapping with high robustness and low drift. *The Journal of Field Robotics*, 35(8), 1242–1264. Available from: https://doi.org/10.1002/rob.21809

Zhang, J., Yao, Y. & Deng, B. (2022) Fast and robust iterative closest point. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7), 3450–3466. Available from: https://doi.org/10.1109/TPAMI.2021.3054619

Zhang, L., Camurri, M., Wisth, D. & Fallon, M. (2021) Multi-camera lidar inertial extension to the newer college dataset. *arXiv Prepr. arXiv2112.08854*. Available from: https://doi.org/10.48550/arXiv.2112.08854

Zhang, Z. & Scaramuzza, D. (2018) A tutorial on quantitative trajectory evaluation for visual(-inertial) odometry. In: *2018 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 7244–7251. Available from: https://doi.org/10.1109/IROS.2018.8593941

Zhao, S., Zhang, H., Wang, P., Nogueira, L. & Scherer, S. (2021) Super odometry: IMU-centric LiDAR-visual-inertial estimator for challenging environments. In: *2021 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 8729–8736. Available from: https://doi.org/10.1109/IROS51168.2021.9635862

Zhao, Y., Zhao, H., Radanovic, M. & Khoshelham, K. (2022) A unified framework for automated registration of point clouds, mesh surfaces and 3D models by using planar surfaces. *The Photogrammetric Record*, 37(180), 366–384. Available from: https://doi.org/10.1111/phor.12428

Zheng, C., Lyu, Y., Li, M. & Zhang, Z. (2020) Lodonet: a deep neural network with 2d keypoint matching for 3d lidar odometry estimation. In: *Proceedings of the 28th ACM international conference on multimedia*, pp. 2391–2399. Available from: https://doi.org/10.1145/3394171.3413771

Zheng, C., Zhu, Q., Xu, W., Liu, X., Guo, Q. & Zhang, F. (2022) FAST-LIVO: Fast and tightly-coupled sparse-direct LiDAR-inertial-visual odometry. In: *2022 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 4003–4009. Available from: https://doi.org/10.1109/IROS47612.2022.9981107

Zheng, X. & Zhu, J. (2021) Efficient LiDAR Odometry for autonomous driving. *IEEE Robotics and Automation Letters*, 6(4), 8458–8465. Available from: https://doi.org/10.1109/LRA.2021.3110372

Zheng, X. & Zhu, J. (2022) Effective solid state lidar odometry using continuous-time filter registration. *arXiv Prepr. arXiv2206.08517*. Available from: https://doi.org/10.48550/arXiv.2206.08517

Zhong, S., Qi, Y., Chen, Z., Wu, J., Chen, H. & Liu, M. (2022) Dcl-slam: a distributed collaborative lidar slam framework for a robotic swarm. *arXiv Prepr. arXiv2210.11978*. Available from: https://doi.org/10.48550/arXiv.2210.11978

Zhong, Y. (2009) Intrinsic shape signatures: a shape descriptor for 3D object recognition. In: *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, pp. 689–696. Available from: https://doi.org/10.1109/ICCVW.2009.5457637

Zhou, B., He, Y., Qian, K., Ma, X. & Li, X. (2021) S4-SLAM: a real-time 3D LIDAR SLAM system for ground/watersurface multi-scene outdoor applications. *Autonomous Robots*, 45, 77–98. Available from: https://doi.org/10.1007/s10514-020-09948-3

Zhou, B., Tu, Y., Jin, Z., Xu, C. & Kong, H. (2023) HPPLO-net: unsupervised LiDAR odometry using a hierarchical point-to-plane solver. *IEEE Transactions on Intelligent Vehicles*, 9(1), 2727–2739. Available from: https://doi.org/10.1109/TIV.2023.3288943

Zhou, H., Lv, K., Huang, L. & Ma, X. (2022) Quantum network: security assessment and key management. *IEEE/ACM Transactions on Networking*, 30(3), 1328–1339. Available from: https://doi.org/10.1109/TNET.2021.3136943

Zhou, P., Guo, X., Pei, X. & Chen, C. (2022) T-LOAM: truncated least squares LiDAR-only odometry and mapping in real time. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–13. Available from: https://doi.org/10.1109/TGRS.2021.3083606

Zhou, Y., Gallego, G. & Shen, S. (2021) Event-based stereo visual odometry. *IEEE Transactions on Robotics*, 37(5), 1433–1450. Available from: https://doi.org/10.1109/TRO.2021.3062252

Zhu, F., Ren, Y., Kong, F., Wu, H., Liang, S., Chen, N. et al. (2023) Swarm-LIO: decentralized swarm LiDAR-inertial odometry. In: *2023 IEEE international conference on robotics and automation (ICRA)*, pp. 3254–3260. Available from: https://doi.org/10.1109/ICRA48891.2023.10161355

Zhu, Y., Zheng, C., Yuan, C., Huang, X. & Hong, X. (2021) Camvox: a low-cost and accurate lidar-assisted visual slam system. In: *2021 IEEE international conference on robotics and automation (ICRA)*, pp. 5049–5055. Available from: https://doi.org/10.1109/ICRA48506.2021.9561149

Zihao Zhu, A., Atanasov, N. & Daniilidis, K. (2017) Event-based visual inertial odometry. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5391–5399. Available from: https://doi.org/10.1109/CVPR.2017.616

Zuo, X., Geneva, P., Lee, W., Liu, Y. & Huang, G. (2019) LIC-fusion: LiDAR-inertial-camera Odometry. In: *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 5848–5854. Available from: https://doi.org/10.1109/IROS40897.2019.8967746

Zuo, X., Yang, Y., Geneva, P., Lv, J., Liu, Y., Huang, G. et al. (2020) Lic-fusion 2.0: Lidar-inertial-camera odometry with sliding-window plane-feature tracking. In: *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 5112–5119. Available from: https://doi.org/10.1109/IROS45743.2020.9340704

How to cite this article: Zhang, Y., Shi, P. & Li, J. (2024) 3D LiDAR SLAM: A survey. *The Photogrammetric Record*, 39, 457–517. Available from: https://doi.org/10.1111/phor.12497