Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 13
AMMSAC 2026

Analysis of SLAM-Based Path Planning Strategies for Mobile Robots

Qianyi Zhang *

College of Information Engineering, Zhejiang University of Technology, Hangzhou, 310000, China

* Corresponding Author Email: IanZhang.Robotic@outlook.com

Abstract. Simultaneous localization and mapping, (SLAM), is universally utilized in mobile robot systems to enable mobile robots to achieve accurate perception and decision-making. Aiming at reviewing the development and research progress of SLAM technology, this review introduces academic research findings on the front-end, back-end, and path planning strategies of SLAM systems focusing on fundamental principles of SLAM systems. Followed by discussion focusing on technology such as semantic maps, deep learning, etc., which are regarded as current research hotspot, the review brings out a brief outlook on future development of SLAM system. Through analysis of current SLAM technology, the review aims to not only summarize challenges current SLAM systems face, but also to offer some guidance. The review, which begins with its discussion from basic process of SLAM technology, highlights the fundamental role SLAM systems play in precise navigation and rapid path planning of the mobile robot systems. Systematically summarizing existing solutions in complex scenarios, it provides a reference for future research and engineering applications.

Keywords: Mobile Robot; SLAM technology; path planning.

1. Introduction

Robots, a milestone in intelligent machinery, are widely applied in manufacturing, healthcare, deep-see/deep-space exploration, and military defense, etc. Compared to those working at a fixed position, mobile robots demonstrate more adaptability in dealing with complex tasks and human-robot interaction. SLAM technology is widely utilized in mobile robot systems to meet their needs for perceiving real-time environments and making rapid decisions. The SLAM technology aims to form the map of the environment rapidly to localized system position, enabling robot systems to navigate autonomously. SLAM technology does not rely on existing maps, thus providing mobile robot systems with greater versatility. Therefore, research on SLAM systems is the foundation of manufacturing, optimizing and applying mobile robots.

Given the cruciality of SLAM, however, traditional SLAM systems, such as Visual SLAM(V-SLAM), LiDAR SLAM etc., exhibit disadvantages in different aspects due to limitations in sensors and algorithms. For instance, V-SLAM systems have high sensitivity to changes in light, while the field of view of LiDAR SLAM systems can be easily narrowed. To solve these, SLAM systems with more efficiency and adaptability emerge.

Focusing on SLAM technology, the review begins with principles of SLAM and moves on to the latest progress on vital technology of both front-end and back-end. As for the path planning strategies, with the concern of real-time performances, the review concentrates on a combination of path planning and technologies such as Deep Learning, semantic maps etc. Ending the review is the outlook based on current SLAM technology, which offers a few potential directions for future development of SLAM system.

2. Basic Principles and Development of SLAM Systems

SLAM technology is aimed at enabling mobile robot systems to navigate autonomously. Different types of SLAM systems have similar principles: acquiring information of surroundings with certain sensors(camera, laser radar etc.) as input. Processing input data with front-end processes, which mainly include feature extraction and feature matching, before map formation, followed by back-end optimization including map optimization, loop closure etc. Finally, a well-formed and optimized map is presented, based on which mobile robot systems choose path planning algorithms and displace accurately. Following paragraphs will cover feature extraction, feature matching, map optimization loop closure and path planning algorithms and their pros and cons.(**Fig. 1** illustrates how a SLAM system works)

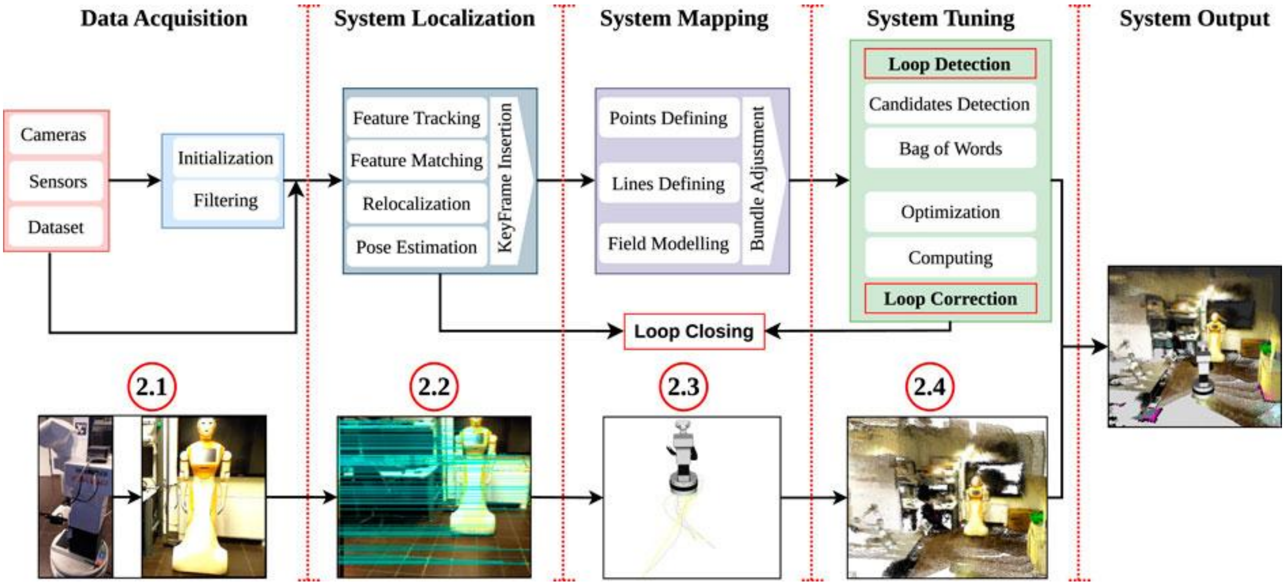

Fig. 1 Visual SLAM architecture (data acquisition, localization, mapping, loop closure) [1]

The V-SLAM acquires information with cameras, including monocular cameras, stereo cameras, RGB-D camera, and event camera etc. The stereo camera captures the environment's depth of field which contributes to the map accuracy, while a monocular camera cannot. The RGB-D camera captures both depth of field and color information. As for the event camera, which switches sampling frequency based on requirements, it performs more accurately and robustly capturing high-speed objects in the environment with changing natural light.

The LiDAR SLAM system has its pre-mapping process mainly involving odometry(scan-to-map), loop closure, and global optimization, from which the odometry process has garnered significant attention from the academic community due to crucial role it plays in estimating the pose between consecutive frames. There are two mainstream approaches to odometry, the scan-to-scan and scan-to-map, with their advantages and disadvantages. Scan-to-scan offers greater real-time capability and cost-effectiveness by directly aligning the scan point clouds of consecutive frames, however, introduces higher error and drift. While the scan-to-map approach directly aligns point clouds with existing maps. Its effective error control makes it more robust than previous methods, but it incurs significant computational overhead due to the need for large-scale map-point cloud matching.

With the aim of better map formation and enhancing system accuracy, multi-sensor fusion SLAM system combines camera, LiDAR, IMU etc. Though challenges in robustness and cross-modal feature matching still exist.

As can be seen, challenges in V-SLAM, LiDAR SLAM and multi-sensor fusion SLAM share common characteristics. Introducing DL, neural network and semantic map etc. is expected to offer solutions from various aspects.

3. Vital Technology of SLAM Systems

The review divided the technology of the SLAM system into front-end technologies and back-end technologies. The front-end technologies, which mainly include feature extraction and feature matching, focus on feature processing of captured images to provide an accurate, efficient description of surroundings for map formation. While the back-end technologies, which mainly involve map optimization and loop closure, provide feedback, filtering, and optimization for SLAM system workflow.

3.1. Feature Extraction and Matching

The feature extraction process and feature matching process, the vital process to form keyframes, not only contribute to pose estimation of the robot systems but also relate to map formation and loop closure.

Sarlin et al. developed a method to find matching features while removing non-matching points with neural network named SuperGlue. SuperGlue neural network, through end-to-end training, possesses the ability to simultaneously infer the mapping relationship between scenes and features, demonstrating superior performance in complex environments [2,3]. Lindenberger et al. improved SuperGlue into an algorithm with higher memory and computational efficiency, greater matching accuracy and lower training difficulty, which they named LightGlue. The LightGlue algorithm offers a new approach in feature extraction and feature matching [4]. Also concentrating on end-to-end training, Tyszkiewicz et al. brought out the DISK method based on DL. The DISK method aims at solving the problem that sparse key points selection and matching exhibit Discreteness, causing end-to-end training difficult to apply. The DISC method demonstrates high accuracy and excellent convergence, offering another well-performed example of end-to-end training [5].

Sun et al. proposed LoFTR, a novel approach in feature matching that distinguishes itself from traditional feature matching methods based on “feature detection, description, and matching” process. To be specific, for feature matching, LoFTR first performs coarse pixel-level dense matching before refining high-quality matches. For feature description, LoFTR utilizes Transformers to generate feature descriptors that simultaneously depend on two images. LoFTR demonstrates superior adaptability to texture-sparse regions compared to traditional models [6]. Similarly facing the issue of poor performance of traditional approaches in in sparsely textured scenes, Wang et al. proposed MatchFormer. MatchFormer, based on Transformer layers, alternately employs self-attention and cross-attention mechanisms for feature extraction and matching, ensuring processing speed, accuracy and robustness [7].

Approaches of feature extraction and feature matching can be divided into detector-based approach and detector-based approach. In addition to approaches above, others exist such as SuperPoint which simultaneously performs feature detection and descriptor learning; R2D2, which relies on keypoints that combine saliency and repeatability.

3.2. Graph Optimization and Loop Closure

The back-end technologies include map optimization (perform nonlinear least-squares optimization on the factor graph to solve for the most consistent global state) and loop closure (detect whether the robot has revisited a previously visited location to correct cumulative drift).

Kaess et al. introduced Bayes Tree into factor graphs, based on which a new sparse nonlinear incremental optimization algorithm, iSAM2, was developed. Compared to traditional optimization algorithms, iSAM2 achieves higher efficiency through incremental variable reordering and streaming linearization. It has been demonstrated that iSAM2 has high map optimization quality and efficiency. As an incremental optimization paradigm, the iSAM2 algorithm is better suited for handling large-scale tasks [8]. Based on Bayers Tree, Zhang et al. proposed the Multi-Root Bayesian Tree (MRBT) to address multi-robot scenarios. MRBT provides a more refined interpretation of sparsity and information flow in SLAM inference problems. Consequently, the graph optimization algorithm MRiSAM2, proposed based on MRBT, achieves more efficient information propagation between root nodes and supports multi-robot inference. In processing multi-robot problems, MR-iSAM2 performed more efficiently compared to iSAM2 algorithm [9]. The ORB-SLAM2, proposed by Mur-Artal et al., combines map optimization with loop closure. To be specific, ORB-SLAM2 incorporates loop closure results into the global pose graph optimization. Experimental results under monocular, stereo, and RGB-D inputs demonstrate that ORB-SLAM2 achieves high accuracy and robustness across multiple indoor and outdoor datasets [10].

Kim et al. proposed Scan Context++, a method for position recognition based on structural appearance. Scan Context++ enables LiDAR point cloud position recognition and closed-loop control. Scan Context++ exhibits robustness against translation and rotation variations. Tests on datasets demonstrate that Scan Context++ achieves both high loop detection rates and low error rates in complex scenes [11]. Xiang et al. proposed another loop detection scheme called FastLCD based on synthetic descriptors and machine learning. FastLCD inputs the specific values of descriptors into the machine learning model to determine whether the laser scan is closed loop. Machine learning enables FastLCD to achieve higher recall rate and better precision [12]. Liu et al. combined a high-frequency odometry based on scan-to-submap approach with lightweight loop detection. They employs lightweight global descriptors for loop candidate retrieval and performs incremental optimization based on sparse pose graphs to correct drift. Experiments on public datasets across diverse scenarios demonstrate that Liu et al. has achieved both high efficiency and precision by integrating scan-to-submap with loop detection [13].

Approaches of map optimization and loop closure mentioned above focusing on specific scenarios, offer considerable improvements over traditional approaches. The academic attention has focused on incremental optimization in map optimization and the integrated treatment of graph optimization and closed-loop control. In closed-loop control, the emphasis lies on multimodal closed-loop detection and closed-loop detection based on global descriptors. What is worth noting is the commonality between research on map optimization and loop closure, which is, most of the research focus on engineering application feasibility and efficiency, trying to come up with new solutions in new practical scenarios.

4. Path Planning Strategies Based on SLAM

When map formation is completed, based on the existing map, mobile robots start to find its path. The path planning process can be divided into two types, global planning and local planning. Former one indicates path planning based on global map and generates optimal path connecting starting point and the destination. The latter one based one system real-time location to operate obstacle avoidance. To achieve these with dynamically changing map from SLAM systems, it is significantly important to ensure real-time performance. Both global and local real-time performance relies on back-end optimization to ensure map consistency. Apart from real-time planning performance, when it comes to planning the path, solutions based on deep learning and semantic maps offer superior precision and robustness, making them a hot topic in academic research.

4.1. Global/Local Planning and Real-time Challenges

The real-time performance of global planning is significantly determined by the process of back-end optimization. To be specific, it can be concluded that map optimization and closed-loop detection methods are prerequisites for achieving excellent real-time global planning. Hu et al. proposed an efficient path planning algorithm for mobile robots based on LiDAR SLAM and an optimized visibility graph. By combining bidirectional A* path search with the minimum construction algorithm to optimize the visibility graph, it achieves high efficiency in static environments global planning [14]. Blöchliger et al. brought out the Topomap framework, which converts sparse feature maps generated by V-SLAM systems into three-dimensional topological maps. The Topomap framework demonstrates advantages in dynamic map updates and navigation in large spaces. The rapid performance of topological maps in global planning achieves excellent real-time performance according to Blöchliger et al.[15]. Wang et al. focus on factory safety inspection scenarios, combining multi-line LiDAR with Visual SLAM systems. They utilize the G2O (General Graph Optimization) library for back-end constraint solving to ensure mapping consistency. AS for path planning, they employ A* for global planning and TEB (Timed Elastic Band). Data demonstrates that the approach performs well in both global route planning and local obstacle avoidance, meeting the demands of complex application scenarios [16].

Global and local real-time performance characterizes the capabilities of path planning algorithms. For the following path planning algorithm to be described, real-time capability is an essential requirement for achieving efficiency and accuracy.

4.2. Path planning Combined with DL and Semantic Maps

Path planning based on SLAM requires efficiency, accuracy and robustness, resulting academic research shifted the focus from isolated algorithms optimization to integration of deep learning and semantic maps.

Wang et al. proposed the Mask-RCNN model that integrates DL with semantic segmentation. Based on RGB-D cameras, Mask-RCNN constructs stable semantic maps through semantic exclusion and generates high-precision maps via graph optimization. Based on Mask-RCNN model, the trajectory drift of subsequent path planning is significantly reduced [16]. Alqobali et al. integrated LiDAR SLAM and V-SLAM to construct semantic maps for indoor mobile robot navigation tasks. The fused semantic map provides multi-level cost information, enabling robots to prioritize the specific safe paths (such as routes avoiding crowds) as its top choice. Their research establishes a paradigm for applying semantic maps in SLAM path planning [17]. Khan et al. proposed an algorithm called RS-DDPG based on deep reinforcement learning. By integrating RS-DDPG with semantic segmentation in SLAM systems, they effectively addressed the susceptibility of V-SLAM to disturbances in dynamic scenes. Experiments demonstrated that the RS-DDPG significantly improved both algorithmic accuracy and robustness of traditional solutions[18]. Li et al. introduced an enhanced DeepLabV3+ semantic segmentation model and optical flow (LK Optical Flow) for feature matching. reducing odometry errors in dynamic scenes. In the research, the semantic segmentation model improved SLAM map localization accuracy and large scale of odometry errors in dynamic scenes had been reduced. A new SLAM technology proposed by Liu et al. supports the local planner (DWA), enabling stable local planning [19]. Another new SLAM system, which combines Visual SLAM and DL, proposed by Sun et al. aims at construct object-oriented semantic maps. This SLAM system performs simultaneously in localization, mapping, and object detection, enabling its application in multi-object path planning tasks. Also, the SLAM model by the Sun team performs with less absolute positioning errors and robust path planning stability [20].

Semantic maps are commonly used in feature extraction to generate cloud maps, while DL can be applied to feature matching, loop closure and map optimization, as well as path planning algorithms. Research above introduces DL and semantic maps into SLAM systems to better optimize path planning. Apart from attaching attention to one technology, there are more research integrating DL with semantic maps. Integrating DL with semantic maps represents a novel and effective research direction for enhancing the efficiency of SLAM systems.

5. Outlooks

Rather than focusing on improvement from an individual technology breakthrough, the development of the SLAM system seeks its future in technology integration. Composite sensors have their natural advantage over applying only one kind of sensor. For instance, V-SLAM, based on camera as a sensor, performed poorly in an environment with varying light, while combing laser and other sensors worked robustly under changing light with camera can solve that from a large scale. Thus, composite sensors offer the system a high data acquisition ability as well as robustness.

Semantic augmentation and semantic maps, which add semantic labels on geometric maps, enable robot "understand" information behind. With the semantic information covered in the map, the SLAM system can effectively mitigate the impact of environmental interference while reducing errors in pose estimation and map formation. The academic research has attached lot attention to semantic augmentation and semantic maps. It is predicted that lager scale of research related to semantic augmentation and semantic maps is expected to broaden the engineering application scenarios of SLAM technology.

6. Conclusion

This review describes the latest developments in key front-end and back-end technologies for SLAM systems. The review is convinced that further development of path planning based on SLAM systems, as concluded in the review, is from single sensor to composite sensor and attaches more importance to the assistance that semantic augmentation, semantic maps, and DL, neural networks can offer to decision making. As mobile robots are increasingly gaining prominence, the SLAM system can be described as self-evidently important. Provided by this review is a thorough introduction of SLAM systems and guidance for further research through an overview of the current development, challenges, and prospects of SLAM systems.

References

[1] Al-Tawil B, Hempel T, Abdelrahman A, et al. A review of visual SLAM for robotics: Evolution, properties, and future applications. Frontiers in Robotics and AI, 2024, 11: 1347985.

[2] Zhang Y, Shi P, Li J. 3D LiDAR SLAM: A survey. The Photogrammetric Record, 2024, 39(186): 457–517.

[3] Sarlin P-E, DeTone D, Malisiewicz T, et al. SuperGlue: Learning feature matching with graph neural networks. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 4938–4947.

[4] Lindenberger P, Sarlin P-E, Rocco I, et al. LightGlue: Local feature matching at light speed. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 1361–1371.

[5] Tyszkiewicz M, Fua P, Trulls E. DISK: Learning local features with policy gradient. Advances in Neural Information Processing Systems (NeurIPS), 2020, 33: 14254–14265.

[6] Sun J, Shen Z, Wang Y, et al. LoFTR: Detector-free local feature matching with transformers. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 8922–8931.

[7] Wang T, Lin J, Li S. MatchFormer: Interleaving attention in transformers for feature matching. Proceedings of the British Machine Vision Conference (BMVC), 2022.

[8] Kaess M, Johannsson H, Roberts R, et al. iSAM2: Incremental smoothing and mapping using the Bayes tree. The International Journal of Robotics Research, 2012, 31(2): 216–235.

[9] Zhang Y, Huang J, Wang H, et al. MR-iSAM2: Incremental smoothing and mapping with multi-root Bayes tree for multi-robot SLAM. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2021: 5090–5096.

[10] Mur-Artal R, Tardós J D. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Transactions on Robotics, 2017, 33(5): 1255–1262.

[11] Kim G, Kim A, Kim A. Scan Context++: Structural place recognition robust to rotation and lateral variations in urban environments. IEEE Transactions on Robotics, 2021, 37(6): 1858–1871.

[12] Xiang H, Shi W, Fan W, et al. FastLCD: A fast and compact loop closure detection approach using 3D point cloud for indoor mobile mapping. International Journal of Applied Earth Observation and Geoinformation, 2021, 102: 102409.

[13] Liu Y, Zhang W, Wu Z, et al. Real-Time Lidar Odometry and Mapping with Loop Closure. Sensors, 2022, 22(3): 1046.

[14] Hu Y, Xie F, Yang J, et al. Efficient path planning algorithm based on laser SLAM and an optimized visibility graph for robots. Remote Sensing, 2024, 16(16): 2938.

[15] Blöchliger F, Fehr M, Dymczyk M, et al. Topomap: Topological mapping and navigation based on visual SLAM maps. arXiv preprint arXiv:1709.05533, 2017.

[16] Wang Z, Zhang Q, Li J, et al. A computationally efficient semantic SLAM solution for dynamic scenes. Remote Sensing, 2019, 11(11): 1363.

[17] Alqobali R, Alnasser R, Rashidi A, et al. A real-time semantic map production system for indoor robot navigation. Sensors, 2024, 24(20): 6691.

[18] Khan S, Niaz A, Dou Y, et al. Deep reinforcement learning and robust SLAM based robotic control algorithm for self-driving path optimization. Frontiers in Neurorobotics, 2025, 18: 1428358.

[19] Li D, Shi X, Long Q, et al. DXSLAM: A robust and efficient visual SLAM system with deep features. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.

[20] Sun R-H, Zhao X, Wu C-D, et al. Research on mobile robot navigation method based on semantic information. Sensors, 2024, 24(13): 4341.