Proceedings of the 4th International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/2026.33872

A Survey of the Evolution of SLAM Techniques for Autonomous Robots (2016–2026)

Jiapeng Zhou

Faculty of Science and Engineering, University of Nottingham, Ningbo, China
scyjz22@nottingham.edu.cn

Abstract. Simultaneous Localization and Mapping (SLAM) is essential for autonomous robots to navigate in unknown environments. SLAM has been applied in diverse scenarios and integrated into a variety of optimization methods and algorithms. Between 2016 and 2026, due to improvements in sensing technologies, SLAM has significantly developed, updated with optimization methods and machine learning. In this essay, the development of key SLAM methods visual, LiDAR-based and visual systems are considered. It emphasizes the change of the old filtering methods to graph-based optimization and enhanced scalability and accuracy of estimation. The essay also talks about the latest trends in incorporating deep learning and semantic understanding. Comparison of various SLAM methods is given in terms of accuracy, robustness and computing cost. Although some significant progress has been made in the past, there are still challenges like dynamic environments, sensor uncertainty and real-time computation. The future work could anticipate working on multi-sensor fusion and learning based SLAM systems.

Keywords: SLAM, LiDAR, IMU

1. Introduction

One of the underlying issues in autonomous robotics is Simultaneous Localization and Mapping (SLAM). SLAM allows a robot to create a map of the unknown world and approximate its location on the map. SLAM is vital in a wide range of applications, such as autonomous vehicles, drones and service robots. Significant improvements in sensing technologies, computational forces and algorithm design have occurred over the last decade that have turned SLAM into a practical solution rather than an abstract idea.

Over the period of 2016-2026, the research on SLAM has been evolving rapidly. In the early days of SLAM, probabilistic filtering methods were used as the main basis. Subsequently, other related techniques have been revised following the rising need to have a high level of dependability and real-time autonomy. The modern methodologies are primarily founded on optimization-based systems, multi- sensors fusion, and machine learning strategies. The developments have enhanced the precision, strength and scalability of SLAM systems. As a result, robot performance has risen and been implemented in more sophisticated settings.

This essay outlines an overview of the development of SLAM methods to autonomous robots within the time span of 2016 to 2026. The survey is limited to the classification according to sensor

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.33872

modalities, such as visual SLAM, LiDAR-based SLAM, and visual–inertial SLAM. It has compared and analyzed in order to assess various methods. The primary issues and emerging research trends are discussed.

2. Background of SLAM

The SLAM problem is a two-task problem that is closely coupled: localization and mapping. Localization can be said to be the estimation of the pose of the robot (position and orientation), and mapping can be said to be the construction of a representation of the surrounding environment, localization and mapping are mutually reliant activities, where map construction cannot be done without accurate pose estimation and vice versa.

The general parts of a SLAM system are a few key components namely sensor data acquisition, feature extraction or scan matching, state estimation, loop closure detection, and map optimization. Cameras, LiDAR sensors, and inertial measurement unit (IMUs) are among the commonly used sensors in SLAM. SLAM systems can take one of the following forms depending on the sensor configuration and include the following visual SLAM, LiDAR SLAM and multi sensor SLAM.

Probabilistic filters that were used in past research to create SLAM algorithms include the Extended Kalman Filter (EKF) and particle filters. These algorithms approximate the combined probability distribution of the map features and robot pose. They have scalability issues and linearization errors, however. In order to address these problems, current SLAM systems are likely to use graph-based optimization algorithms, where the optimization problem is formulated as a nonlinear least-squares optimization of a pose graph [1].

3. Evolution of SLAM techniques (2016–2026)

3.1. Transition from filtering to optimization

Until 2016, numerous SLAM systems were recursive filtering based. These methods, however, had a weakness of scaling due to the increased state space and cumulative linearization errors. With the introduction of graph-based SLAM, a tremendous jump was achieved, and the problem could then be expressed as a global optimization problem [1]. This update was made to make it possible to improve accuracy and consistency, which might lead to better performance in large-scale settings.

From 2016 to 2020, it became the popular paradigm to use optimization-based SLAM. Both the combination of bundle adjustment and pose graph optimization had sparked substantial enhancement in the trajectory estimation and loop closing performance.

3.2. Visual SLAM: feature-based to learning-based systems

Visual SLAM was associated with tremendous growth as camera sensors could be installed. The first systems such as ORB-SLAM2 were feature-based that identified and matched salient points across frames. These techniques are effective in organized environments where there are adequate features of visualization [2].

However, the reliance on feature extraction shows the weaknesses of visual SLAM. Feature matching cannot be relied upon if conditions have bad texture or change in lighting. The match problems were solved by introducing direct methods, which match features based on pixel intensity values. They are stronger in some cases, although it requires additional resources to compute [2, 3].

DOI: 10.54254/2755-2721/2026.33872

Since 2020, deep learning has started to affect visual SLAM in several features: Depth estimation, feature extraction and place recognition are increasingly completed with neural networks. These methods enhance stability in complicated environments, although they have introduced new issues, including generalization and computational cost [4].

3.3. LiDAR SLAM: high precision and large-scale mapping

LiDAR-SLAM has been a significant technology in the uses of high-precision such as autonomous driving. Such algorithms as LOAM decouple the odometry/mapping tasks, improving efficient real-time operation.

LiDAR can offer actual distance measurements, which makes it less susceptible to the lighting conditions than cameras do [5]. And this enables LiDAR SLAM to work reliably in outdoor and low-visibility conditions. Nonetheless, large point clouds are still computationally expensive to process, and sensor cost is still an issue.

Recent advancements have been made concerning enhancing robustness in feature-sparse settings and minimizing computing costs. Some of the techniques that have been suggested and tested include scan matching optimization, and hybrid techniques that combine LiDAR and other sensors [5].

Scan matching compares point clouds of successive frames to determine the motion of the robot, which is donated as a vital element of LiDAR SLAM. Two common algorithms are: Iterative Closest Point (ICP) and Normal Distributions Transform (NDT). ICP Reduces the distance between similar points and NDT Represents point clouds as probability distributions.

These techniques allow proper motion estimation. Nevertheless, they are computationally intensive, especially with dense point clouds.

Newer methods, like LOAM (LiDAR Odometry and Mapping), where odometry and maps are split into two independent processes. The approach enhances efficiency since it does not only carry out fast motion estimation but also ensures high-quality mapping in the background [6].

3.4. Visual–inertial SLAM: enhancing robustness through fusion

Visual inertial SLAM takes into account both camera and IMU data to enhance motion estimation. These systems are able to coordinate visual measurements with inertial measurements to provide the correct tracking when there is rapid motion or the momentary loss of visual details.

Graph optimization strategies are used in frameworks like VINS-Mono to achieve proper state estimation and minimize drift [7]. The availability of IMU data gives the high-frequency motion updates further enhancing system stability.

Although they have these benefits, visual-inertial SLAM adds more complexity. As there are several sensors, sensors have to be calibrated and synchronized correctly and failure in the process can lead to high performance failure [7].

3.5. Emerging trends: semantic and learning-based SLAM

SLAM research has been on the rise in popularity with machine learning methods since 2020. Semantic SLAM builds on its previous counterparts by incorporating the concept of object level understanding in the mapping.

Traditional SLAM algorithms assume a fixed environment and make use of low-level geometrical sources like points, edges or planes. The real world is, however, dynamic and complex.

DOI: 10.54254/2755-2721/2026.33872

Moving objects, altering lighting conditions and constrictions. Organizations accumulate complex problems that cannot be adequately tackled by a mere geometric approach.

Semantic SLAM builds on the existing systems by incorporating high level knowledge, including object categories and scene analyzing. Semantic SLAM helps robots to identify and categorize objects rather than viewing the environment as a set of abstract features [8]. This improves the accuracy of mapping and decision making.

To enhance perception and robustness, learning-based SLAM learned data-driven models, in particular, deep neural networks. These strategies will help minimize dependence on hand-crafted functionality and enhance the flexibility to different environments and varying conditions.

Also, neural rendering methods like Neural Radiance Fields (NeRF) are capable of producing high-quality grid, location and coordinates 3D reconstruction, which is a trend in data-driven mapping [9]. These approaches are both promising but they are very intensive in terms of the required computing resources and are currently being developed.

4. Comparison of SLAM techniques

SLAM methods can be measured using a number of performance criteria, such as localization accuracy, resistance to environmental variations, computational cost, and scalability. The three major categories, which include visual SLAM, LiDAR SLAM and visual inertial SLAM provide different trade-offs. Visual SLAM systems are better than the other systems as they cost less in hardware and they can be used to extract rich environmental information. Yet, they are very sensitive to visual conditions. Tracking failure and higher estimation error may occur due to factors like variation of lighting, motion blur, and absence of texture [10]. Monocular systems too are affected by scale ambiguity unless further restrictions are imposed.

The LiDAR-based SLAM is much more accurate in geometry with direct distance measurements making it very useful on large scales and outdoors. In addition, LiDAR systems are robust to lighting conditions, which is also a significant advantage compared to camera-based methods. Moreover, the high cost of LiDAR sensors and the computational burden of processing point clouds also limit their accessibility.

Visual–inertial SLAM offers a compromise by combining complementary sensor modes. The integration of IMU data improves motion estimation and reduces drift, especially during fast movements or pose change [11]. However, this approach requires an accurate sensor calibration and requires additional sources of error intrigued by IMU noise.

As shown in Table 1, quantitative evaluation of SLAM systems often relies on metrics such as Absolute Trajectory Error (ATE) and Relative Pose Error (RPE). Benchmark studies indicate that LiDAR SLAM generally achieves the highest accuracy, while visual SLAM offers lower computational requirements [12]. Visual-inertial SLAM trades off between these factors, offering better robustness, at reduced cost compared to LiDAR systems.

Table 1. Comparison of SLAM techniques

Method	Sensor type	Accuracy	Robustness	Computational cost	Advantages	Limitations
Visual SLAM	Camera	Medium-High	Sensitive to lighting conditions	Medium	Low cost, semantic richness	Illumination sensitivity, motion blur
LiDAR SLAM	LiDAR	High	Strong in diverse conditions	High	High precision, reliable depth	High cost, computationally intensive

DOI: 10.54254/2755-2721/2026.33872

Table 1. (continued)

Visual -Inertia SLAM	Camera+ IMU	High	Robust to motion	Medium-High	Reduced drift, improved stability	Calibration complexity, sensor noise

5. Challenges in SLAM

Despite all the achievements, SLAM research still has some problems. Some difficulties include the ability to handle dynamic environments where dynamic objects may cause differences in the map. Most of the SLAM systems operate on the assumption of a fixed world which could limit their applicability to actual conditions in the real static world.

Sensor noise and uncertainty also pose a challenge to the accuracy of the state estimation. Measurement error can drift with time. This is amenable to loop closure detection, but even in large scale contexts, detection remains a challenge.

Also, there is another significant issue which is the computational efficiency. Real-time applications on resource-constrained systems, including high-resolution images or thick point clouds, demand a lot of computation.

6. Future trends

The area of the future SLAM research is believed to be the enhancement of robustness, scalability, and adaptability. Multi-sensor fusion methods will remain important in the process of synchronizing complementary sensing modalities. In the meantime, Learning-based SLAM approaches are more likely to be accepted. Such strategies will help towards better perception and real time decision making.

The intelligent interaction between robots in their working environment will be possible with the help of semantic SLAM and scene understanding. Furthermore, further research, including Uncertainty Estimation and Optimizing for Resource-Constrained Platforms, will allow real-time SLAM to be used in more complex situations [13].

7. Conclusion

In this paper, the development of SLAM methods to autonomous robots from 2016 to 2026 has been surveyed. The switching between classical probabilistic algorithms to modern optimization-based and learning-driven algorithms have enhanced performance of the SLAM systems. All types of methods, visual, LiDAR, and visual-inertial SLAM, have their own benefits, and they are optimally applied to different situations.

In spite of these developments, issues of dynamism, computational efficiency and sensor constraints are still present. Multi-sensor fusion, semantic understanding and learning-based approaches are likely to be considered in future research to improve SLAM in new ways.

References

[1] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, J.J. Leonard, Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. IEEE Trans. Robot. 32, 1309–1332 (2017)

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

DOI: 10.54254/2755-2721/2026.33872

[3] J. Engel, V. Koltun, D. Cremers, Direct sparse odometry. IEEE Trans. Pattern Anal. Mach. Intell. 40, 611–625 (2017)
[4] A. Beghdadi, M. Mallem, A comprehensive overview of dynamic visual SLAM and deep learning: Concepts, methods and challenges. Mach. Vis. Appl. 33, 54 (2022)
[5] M.U. Khan, S.A.A. Zaidi, A. Ishtiaq, S.U.R. Bukhari, S. Samer, A. Farman, A comparative survey of LiDAR-SLAM and LiDAR-based sensor technologies, in Proc. MAJICC 2021, 1–8 (2021)
[6] H. Wang, C. Wang, C.L. Chen, L. Xie, F-LOAM: Fast LiDAR odometry and mapping, in Proc. IEEE/RSJ IROS, 4390–4396 (2021)
[7] T. Qin, P. Li, S. Shen, VINS-Mono: A robust and versatile monocular visual-inertial state estimator. IEEE Trans. Robot. 34, 1004–1020 (2018)
[8] S. Han, Z. Xi, Dynamic scene semantics SLAM based on semantic segmentation. IEEE Access 8, 43563–43570 (2020)
[9] B. Mildenhall, P.P. Srinivasan, M. Tancik, J.T. Barron, R. Ramamoorthi, R. Ng, NeRF: Representing scenes as neural radiance fields for view synthesis. Commun. ACM 65, 99–106 (2021)
[10] A. Pretto, E. Menegatti, M. Bennewitz, W. Burgard, E. Pagello, A visual odometry framework robust to motion blur, in Proc. IEEE ICRA, 2250–2257 (2009)
[11] T. Pereira, C. Viegas, S. Soares, N. Ferreira, Visual and visual–inertial SLAM for UGV navigation in unstructured natural environments: A survey of challenges and deep learning advances. Robotics 15, 35 (2026) https://doi.org/10.3390/robotics15020035
[12] Y.-L. Zhao, Y.-T. Hong, H.-P. Huang, Comprehensive performance evaluation between visual SLAM and LiDAR SLAM for mobile robots: Theories and experiments. Appl. Sci. 14, 3945 (2024) https://doi.org/10.3390/app14093945
[13] C. Galagain, M. Poreba, F. Goulette, Is semantic SLAM ready for embedded systems? A comparative survey. arXiv preprint arXiv: 2505.12384 (2025)