Visual SLAM Methods for Autonomous Driving Vehicles

Rui Xu *

Huijia Private School, Beijing, China

* Corresponding Author Email: 25xurui@huijia.edu.cn

Abstract. Autonomous driving has recently become a burgeoning field poised to revolutionize transportation, with apparent anticipation about its widespread adoption soon. As vehicles equipped with autonomous capabilities become increasingly prevalent, the need for robust navigation systems becomes vital. Simultaneous Localization and Mapping (SLAM) methods have emerged as a critical solution to address the challenges inherent in autonomous driving. By concurrently creating maps of the environment and accurately localizing vehicles, SLAM algorithms enable autonomous vehicles to navigate safely and efficiently in diverse, dynamic, and even GPS-denied environments. This paper aims to elucidate the functionality and principles underpinning SLAM methods, with a particular focus on their application in autonomous driving vehicles. By examining traditional localization methods and their limitations, this paper underscores the pivotal role of SLAM in overcoming these challenges. Furthermore, this paper delves into the advancements in visual SLAM technology and its effectiveness in resolving contemporary issues encountered by autonomous vehicles, such as uncertainties in urban environments. The integration of Convolutional Neural Networks (CNNs) with visual SLAM systems is discussed, showcasing the potential to enhance depth estimation; optical flow, feature correspondence, and camera pose estimation. Despite these advancements, persistent challenges remain, including map robustness, computational requirements, and security considerations. Nevertheless, by leveraging visual SLAM technology, autonomous driving vehicles are poised to navigate complex environments with unprecedented precision, paving the way for a future where transportation is safer, more efficient, and more accessible than ever before.

Keywords: Autonomous vehicle; visual SLAM method; localization; mapping.

1. Introduction

The quest for autonomous driving technology has sparked a remarkable surge of interest in Simultaneous Localization and Mapping (SLAM) algorithms. These algorithms are integral to enabling intelligent mobile robots to autonomously navigate unfamiliar environments. As advancements in this field continue to reshape the landscape of transportation, it becomes imperative to delve into the intricacies of SLAM methodologies and their specific applications.

This paper undertakes a comprehensive examination of SLAM methods, with a primary focus on their functionality, working principles, and their application within autonomous driving vehicles. It starts with a historical exploration of SLAM, tracing its origins back to its inception in 1986 and highlighting its transformative journey over the years. From its initial conceptualization to its contemporary applications in autonomous navigation, SLAM has emerged as a cornerstone technology, facilitating real-time mapping and localization in dynamic environments.

The discussions throughout this paper encapsulate the fundamental components of SLAM algorithms, with a particular emphasis on visual SLAM techniques, including initialization, tracking, and mapping. By leveraging visual and spatial technologies, visual SLAM methods utilize camera-based sensor systems to gather 2D images and construct comprehensive environmental maps. This paper examines the complexities of visual SLAM, elucidating its role in aiding state estimation and reducing estimation errors during navigation.

Furthermore, this paper delves into the challenges faced by autonomous driving vehicles and the pivotal role of SLAM methods in overcoming these obstacles. Ranging from the limitations of GPS-based localization methods to the complexities of urban environments, autonomous vehicles confront a myriad of challenges that demand robust navigation solutions. Additionally, this paper explores the intersection of SLAM technology and autonomous driving, examining advancements in Convolutional Neural Networks (CNNs) and their application in addressing key challenges such as depth estimation, optical flow, and feature correspondence.

The rest of this paper is organized as follows. Section 2 begins by discussing the historical evolution of SLAM in the context of autonomous navigation, followed by an in-depth examination on visual SLAM techniques and their role in aiding state estimation. Section 3 delves into the challenges faced by autonomous driving vehicles and several solutions to these challenges such as advancements in CNNs. Lastly, this paper concludes by highlighting current research trends and delineating future directions in SLAM technology within the realm of autonomous driving.

2. Functionality and working principles of SLAM method

The SLAM algorithm is an algorithm designed to map and locate an object in an unknown environment simultaneously. First proposed in 1986, this method has undergone continuous advancement [1] and now plays a pivotal role in facilitating autonomous robot navigation. By simultaneously generating a map of the environment and accurately localizing the robot, SLAM becomes an indispensable technology for intelligent mobile robots in scenarios where GPS signals are unavailable [1]. However, challenges persist, such as uncertainties in robot movements. Over the past 25 years, numerous studies and surveys have explored various aspects of SLAM, including filtering, application utilization, sensing-based evolution, etc. [1]. Due to intensive research, SLAM continues to evolve and has become a practical and crucial technology in various applications.

Recently, visually specified SLAM technology has emerged to address contemporary challenges faced by modern robots. Visual SLAM technology represents specialized branch of the SLAM method, focusing on visual and spatial technologies It has emerged as a vital tool for addressing contemporary challenges faced by robots Unlike other methodologies, visual SLAM prioritizes the mapping of unknown environment while locating a sensor system within it. Visual SLAM techniques employ one or more cameras within the sensor system, utilizing 2D images as the primary source of information [2]. The algorithm typically consists of three main components: initialization, tracking, and mapping, as shown in Figure 1 [2]. In the context of robotics, mapping is crucial for visualizing landmarks, assisting state estimation and reducing estimation errors during revisits [2]. The mapping process involves addressing three key questions from the robot's perspective: What does the world look like? Where am I? and how can I reach a given location? [2]. The challenge of choosing the most suitable technique based on project constraints can be addressed through mapping, localization, and path-planning tasks.

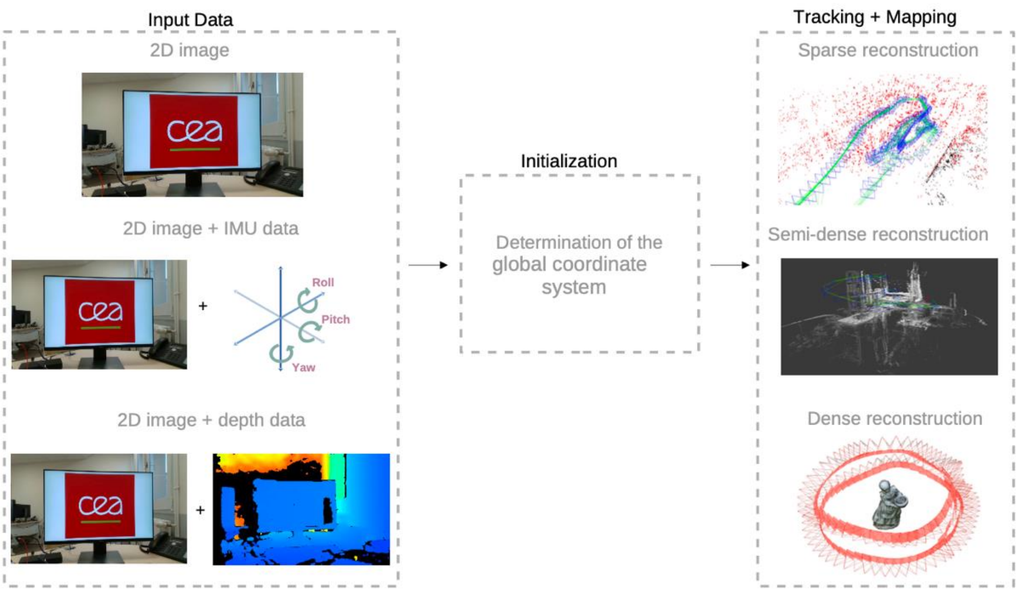

Fig. 1 Three main components of the visual SLAM method [2]

3. Visual SLAM methods for autonomous driving vehicles

In the practical scenarios encountered by autonomous driving vehicles, accurate navigation is crucial, demanding precise representation of the environment and self-state estimation. Traditional localization methods reliant on GPS or RTK encounter limitations due to measurement errors, especially in challenging terrains like tunnels and urban canyons [3]. However, unlike the traditional methods, the SLAM approach enables concurrent real-time pose estimation while constructing an environment map, thereby addressing the majority of challenging scenarios [3]. SLAM methods are typically categorized into Light Detection and Ranging (LIDAR) SLAM and visual SLAM, each offering distinct advantages. LIDAR is a mature solution in autopilot applications, providing a 3D map with a larger Field of View (FOV), albeit with higher costs and longer development cycles. Conversely, despite challenges in urban environments, visual SLAM proves to be cost-effective, lightweight, and rich in information [3].

Nevertheless, current autonomous driving vehicles still face challenges. In this rapidly evolving field, automated driving encompasses two main paradigms: mediated perception and behavior reflex [4]. Mediated perception involves semantic reasoning of the scene to determine driving decisions, while behavior reflex learns end-to-end driving decisions [4]. One of the vital techniques utilized to address these challenges is Convolutional Neural Networks (CNNs).

CNNs serve as crucial tools in computer vision, playing fundamental roles in tasks like image classification and target detection [4]. Depth estimation, a critical aspect of automated driving, has evolved through CNN-based approaches, employing regression losses and regularization terms for supervised learning or unsupervised techniques based on multiple views and photometric error assessment [4]. Many aspects of automated driving have become more advanced due to the development of CNNs. For example, optical flow, a vital aspect of motion detection, has seen significant advancements with CNN-based methods outperforming classical geometry-based approaches [4]. Feature correspondence, another fundamental aspect, can also benefit from CNN-based techniques, such as universal correspondence networks [4]. While CNN-based bundle adjustment solutions are still in their infancy, semantic segmentation, a crucial technique for image partitioning in various applications, has witnessed developments in three subcategories: patch-wise training, end-to-end learning, and multi-scale segmentation [4]. Additionally, camera pose estimation, an essential part of localization in SLAM, has explored CNN-based approaches like Pose Net and methodologies integrating geometrical constraints for superior results [4]. Figure 2 shows the working principle of the Visual SLAM method based on CNNs. Despite concerns regarding map retraining, these opportunities in deep learning offer promising avenues for improving Visual SLAM in the context of automated driving.

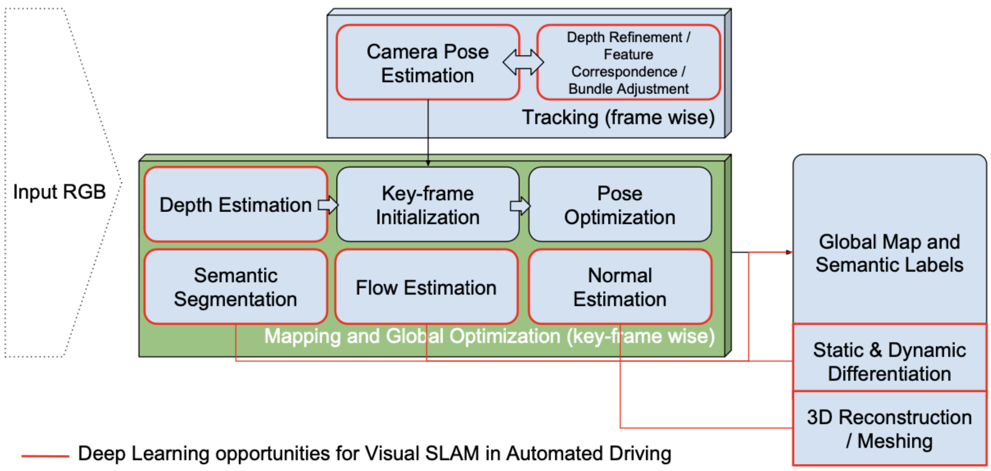

Fig. 2 The diagram shows how Visual SLAM works utilizing depth estimation, pose estimation, etc [4]

On the other hand, while the Visual SLAM method offers solutions to many challenges, it still faces hurdles in urban environments [5]. Current Visual SLAM approaches encounter persistent challenges and limitation, such as maintaining map robustness amidst scene changes, reliance on high-end GPUs for deep learning, handling continuous open-world scenes, and adapting to dynamic changes in the environment [5]. Additionally, there is a demand for algorithmic simplification. The unsolved nature of SLAM is marked by its dependency on the environment, robot characteristics, sensor uncertainties, and performance goals. Therefore, the security of the automated vehicle becomes a concern as well [5].

Autonomous driving safety is a pressing concern as interconnected autonomous vehicle networks rely heavily on digital systems and communication frameworks. These networks operate at two main levels: vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) [5]. Both levels facilitate the exchange of critical information for efficient navigation. However, the complex architecture of these networks makes them vulnerable to a variety of security threats, broadly categorized into passive and active attacks [5]. Passive attacks involve eavesdropping on or analyzing signals to obtain sensitive information without changing the functionality of the system [5]. Although encryption is used to protect data, passive attacks can successfully decrypt the information. Active attacks are more widespread, including man-in-the-middle attacks, denial of service (DoS) attacks, and replay attacks [5]. Vulnerabilities present in controlling various components of self-driving systems are a serious concern, especially considering that the Controller Area Network (CAN) protocols are widely utilized in most self-driving cars [5]. The broadcast nature of CAN, vulnerability to DoS attacks, and lack of proper authentication mechanisms expose potential risks and emphasize the urgent need for continuous research and improvement in autonomous driving safety [5].

In the field of autonomous driving, mapping is the most crucial section. Autonomous driving cars need mapping to adjust directions and plan routes [6]. Therefore, maps become indispensable for ensuring the safe and efficient operation of self-driving vehicles. Besides navigation, maps serve as foundational elements that transcend the limitations of onboard sensors [6]. Unlike human drivers who can adapt to unforeseen circumstances, autonomous vehicles require meticulous and highly detailed maps to guide their decision-making processes [6]. These maps act as trusted references, especially when sensor availability is compromised. The complexity of autonomous driving challenges traditional transportation norms, infrastructure, and urban development, underscoring the indispensable role of maps in integrating this transformative technology successfully.

One distinctive advantage of SLAM maps in autonomous vehicles is their unparalleled capabilities compared to other sensors [6]. With an infinite range, maps enable vehicles to "see" into occluded areas, providing a comprehensive understanding of the environment [6]. Moreover, maps offer a level of reliability that surpasses environmental conditions, making them a failsafe source of redundancy in safety-critical autonomous applications [6]. The highly refined data contained in maps, meticulously processed and verified over hours or days, empowers them to furnish accurate, meaningful, and real-time information. However, potential inaccuracies, inconsistencies, and outdated information pose risks, and the high precision demanded by autonomous driving amplifies the complexity and cost of map production [6].

One way to improve the mapping functionality is to build a more sophisticated visual SLAM system. In the quest for achieving more accurate and computationally efficient visual SLAM systems, one of the solutions is to combine sparse and dense mapping techniques [7]. The need for dense maps is particularly crucial where tasks such as path planning, navigation, and collision avoidance demand detailed and comprehensive environmental representations [7]. The new system in [7] builds upon the foundation of sparse V-SLAM, which utilizes an Extended Kalman Filter (EKF) for real-time pose estimation and landmark mapping. It also addresses the computational challenges inherent in EKF SLAM through the utilization of the conditionally independent sub-mapping method [7], ensuring constant run-time complexity for most scenarios and enhancing the efficiency of the V-SLAM system. The second component involves computing a dense point cloud from stereo images, providing a local map more accurate than simply reconstructed disparity images [7]. A dense reconstruction expressed in local coordinate systems is achieved by employing a Kalman Filter for every pixel in the image, considering the local ego displacement [7]. This crucial innovation, stemming from the separation of the sparse and dense maps, offers a significant advancement in the SLAM estimate.

In recent years, advancements in self-driving car technology have spurred an expansion in the scale of experiments aimed at validating location algorithms. Research teams have conducted extensive trials over vast distances due to the soaring of large-scale datasets and dedicated test platforms [8]. Early experiments in the 1990s focused on highway environments, relying on lane detection algorithms for vehicle location [8]. However, major events like the DARPA Grand Challenge have ushered in a new era, prompting researchers to develop more sophisticated positioning techniques to adapt to diverse terrain conditions [8]. During the 244-kilometer race across the desert landscape in the DARPA Grand Challenge, the positioning strategy evolved to combine GPS with elevation maps generated by laser scanners, enabling vehicles to navigate challenging features like cliffs and rocks [8]. Similarly, the subsequent DARPA Urban Challenge, set in an urban environment, underscored the importance of map reliability, with its Route Network Definition File (RNDF) format providing detailed road topologies and geometry [8].

Furthermore, in 2011, collaborative projects like the Grand Collaborative Driving Challenge (GCDC) explored the potential of fleets to alleviate congestion, demonstrating the benefits of sharing location information among fleet members [8]. Additionally, large-scale experiments such as the VisLab team's 13,000km trip from Italy to China have showcased the feasibility of remote autonomous driving through a combination of lane-keeping algorithms and leader-following systems [8].

These experiments emphasize the critical role of robust positioning algorithms in enabling safe and efficient autonomous navigation across various environments, spanning from highways to city streets. As research advances, addressing challenges such as map maintenance, sensor fusion, and scalability will be crucial to realizing the full potential of autonomous driving technology in practical applications.

4. Conclusion

In conclusion, the evolution of the SLAM) methods has played a pivotal role in propelling autonomous driving technology into a transformative era of transportation. From its inception to its contemporary applications, SLAM has emerged as a cornerstone of autonomous navigation, enabling vehicles to traverse safely and efficiently through dynamic and GPS-denied environments. The integration of visual SLAM techniques, leveraging visual and spatial technologies, has proven instrumental in overcoming challenges encountered by autonomous vehicles, particularly in urban settings. Additionally, the incorporation of CNNs has further enhanced the capabilities of visual SLAM, promising advancements in depth estimation, optical flow analysis, feature correspondence identification, and camera pose estimation.

Despite significant progress made so far, challenges such as maintaining the robustness of maps, making computational requirements more efficient, and addressing security considerations continue to exist. These challenges underscore the need for ongoing research and innovation in the field of SLAM. The inherent complexity of SLAM algorithms is compounded by sensor noise and uncertainty; hence probabilistic methods are necessary to effectively manage these challenges. Moreover, the interdependence between localization accuracy and mapping precision poses inherent difficulties that highlight the necessity for standardized methodologies when comparing results across different SLAM algorithms. Nevertheless, with the continuous advancements in SLAM technology and the relentless pursuit of innovation, self-driving vehicles are poised to revolutionize transportation. This promises a future where safety, efficiency, and accessibility redefine how we navigate our world.

References

[1] Taheri H., Zhao C. X., et al. SLAM; definition and evolution. Engineering Applications of Artificial Intelligence, 2021, 97: 104032.

[2] Barros M. A., Michel M., Moline Y., Corre G., Carrel F., et al. A Comprehensive Survey of Visual SLAM Algorithms. Robotics, 2022, 11(1): 24.

[3] Cheng J., Zhang L., Chen Q., Hu X., Cai J., et al. A review of visual SLAM methods for autonomous driving vehicles. Engineering Applications of Artificial Intelligence, 2022, 114: 104992.

[4] Milz S., Arbeiter G., Witt C., Abdallah B., Yogamani S., et al. Visual SLAM for Automated Driving: Exploring the Applications of Deep Learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2018: 247-257.

[5] Singandhupe A., La H. M. A Review of SLAM Techniques and Security in Autonomous Driving. In Proceedings of the Third IEEE International Conference on Robotic Computing (IRC), Naples, Italy, 2019: 602-607.

[6] Lategahn H., Geiger A., Kitt B. Visual SLAM for autonomous ground vehicles. In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 2011: 1732-1737.

[7] Wong K., Gu Y., Kamijo S., et al. Mapping for Autonomous Driving: Opportunities and Challenges. IEEE Intelligent Transportation Systems Magazine, 2021, 13(1): 91-106.

[8] Bresson G., Alsayed Z., Yu L., et al. Simultaneous localization and mapping: A survey of current trends in autonomous driving. IEEE Transactions on Intelligent Vehicles, 2017, 2(3): 194-220.