Automatic driving technology based on SLAM technology

Haoyu Wei*

School of Energy, Power and Mechanical Engineering, North China Electric Power University, Hebei, China

* Corresponding Author Email: 220231041218@ncepu.edu.cn

Abstract. The method known as simultaneous localization and mapping, or SLAM, is used to determine the movements of sensors in the surroundings as well as the 3D structure of an unidentified area. This method was first proposed in robotics to accomplish autonomous control of robots. Subsequently, SLAM-based applications have expanded significantly, including self-driving automobiles, augmented reality (AR)-based visualization, and computer vision-based online 3D modeling. Then, SLAM-based applications have expanded significantly, including self-driving automobiles, augmented reality (AR)-based visualization, and computer vision-based online 3D modeling. This paper synthesizes some other researchers' experimental process and conclusions, and discusses the practicability and feasibility of SLAM technology in automotive autonomous driving. It includes the introduction of SLAM technology, the application of SLAM technology in automobile autonomous driving and the shortcomings and solutions of SLAM technology. A large number of experimental research results have proved the feasibility of this technology for this field, which also means that SLAM technology can play a huge role in the field of automotive autonomous driving.

Keywords: Automatic driving, SLAM, Augmented reality.

1. Introduction

In recent years, significant progress has been made in autonomous driving and advanced driving assistance technology, especially the rapid development of deep learning and other artificial intelligence technologies, and the perception ability of autonomous vehicles has been significantly improved. However, in the actual driving environment, it is almost impossible to obtain "perfect" perception data due to errors or noise in the process of target recognition, detection and tracking. More importantly, in the actual traffic scenario, there is a strong correlation and interaction between the behaviors of autonomous vehicles and surrounding multiple traffic agents, which makes it difficult to accurately predict the behavioral intentions and future trajectories of surrounding multiple targets. In the highly interactive complex driving environment, establishing a behavior decision system that considers the influence of uncertainty factors such as perception and prediction is one of the main problems that the current autonomous vehicle needs to solve. An accurate map of the environment is necessary for a car to drive itself effectively. The availability of a precise map enables the creation of devices that can operate in difficult environments exclusively utilizing their on-board sensors and without the assistance of an external reference system like GPS [1]. This will allow the car to judge its position relative to the environment in the process of autonomous driving more accurately. The simultaneous localization and mapping (SLAM) problem is often referred to as "learning maps under pose uncertainty." Simultaneous localization and mapping, or SLAM for short, is a method for predicting sensor motion and recreating structure in an unidentified environment. For the community studying autonomous vehicles, the solution to the simultaneous localization and map building (SLAM) problem is essentially the "Holy Grail." Indeed, a robot would be considered "autonomous" if it could be programmed to locate itself in an unfamiliar environment, create a map of it using only relative observations, and then use that map to navigate at the same time. Therefore, the primary benefit of SLAM is that it does not require artificial infrastructures or pre-existing topological knowledge of the surrounding area. In the fields of artificial intelligence and mobile robots, simultaneous localization and mapping, or SLAM, is a crucial issue that deals with localization and mapping in situations where a previous map of the workspace is unavailable. In the fields of computer vision, augmented reality, and robotics, SLAM has been studied in the literature.

It is a fundamental technology for many kinds of applications [2]. The literature offers a number of methods for resolving this issue. Filtering techniques model the problem as an online state estimation, where the map and the car's current position comprise the system's state. As fresh measurements become available, they are added to and improved upon the estimate. This group includes well-known methods including information filters, particle filters, and Kalman and information filters. Typically, the filtering procedures are called "online SLAM methods" to emphasize their incremental nature. On the other hand, smoothing techniques calculate the entire car's trajectory using all of the measurements. These methods, which focus on the so-called full SLAM problem, usually use methods for least-square error minimization [1].

2. Overview of SLAM

The method known as simultaneous localization and mapping, or SLAM, is used to determine the movements of sensors in the surroundings as well as the 3D structure of an unidentified area. In the self-driving technology of cars, this method was first proposed to accomplish autonomous control of cars. Subsequently, SLAM-based applications have expanded significantly, including self-driving automobiles, augmented reality (AR)-based visualization, and computer vision-based online 3D modeling. Numerous sensor types, including laser range sensors, rotary encoders, inertial sensors, GPS, and cameras, were combined into the first SLAM algorithms [2].

The so-called graph-based formulation of the SLAM problem offers an intuitive solution. In order to solve a graph-based SLAM problem, one must first build a graph with nodes representing automotive landmarks and edges encoding sensor measurements that limit the connected poses of the nodes. Naturally, as noise continually affects observations, such limitations may be in conflict with one another. When creating such a graph, the key challenge is figuring out a node layout that maximizes consistency with the measurements.

Lu and Milios introduced the graph-based formulation of the SLAM problem in 1997 [3]. However, the relatively high complexity of solving the error minimization problem using normal techniques delayed the popularization of this formulation by several years. The optimization problem at hand was solved effectively thanks to recent developments in sparse linear algebra and insights into the nature of the SLAM problem. As a result, graph-based SLAM approaches have had a resurgence and are now among the fastest and most accurate methods available.

SLAM solution strategies have advanced rapidly over the last two decades. Numerous SLAM algorithms have been created that estimate a robot's pose while concurrently creating two-dimensional (D) or three-dimensional (3D) maps using a variety of sensors, including laser scanners, RGB cameras, and ultrasonic sensors. It is claimed that the 2D SLAM problem with range finders is considered solved (Li et al., 2016) [4]. Nevertheless, this claim is untrue, as there is always room for development in the 2-D SLAM problem's accuracy, speed performance, and capacity to handle large amounts of data. On the other hand, among the most difficult unsolved issues that still need to be resolved are high-quality, reliable, and three-dimensional visual assaults on unmanned systems like Unmanned Aerial Vehicles (UAVs), self-driving cars, building inspections, surveillance, underwater SLAM, and a host of other issues pertaining mostly to outdoor, unstable, and dynamic environments. Furthermore, every technique may always be made better. SLAM failures are primarily caused by unstable perception or uncertainty. SLAM comprises two primary components: mapping and localization. Prior to the development of SLAM technology, localization and mapping were treated as distinct concepts. It was found by contemporary academics that there is a strong internal dependency between the mapping and the localization. While accurate localization depends on the map, mapping depends on localization. Thus, the question is known as a "Chicken and egg" question. More sophisticated methods previously approached localization and mapping as two independent tasks. Traditional SLAM systems collaboratively estimate the poses and the map, however, later, more sophisticated methods saw localization and mapping as two separate processes, leading to the well-known Parallel Tracking and Mapping (PTAM) (Klein and Murray, 2007) [5].

3. The application of SLAM in experiment

H. Lategahn's article proposes a stereoscopic V-SLAM system that can calculate dense maps more accurately than simply reconstructing parallax images. The system consists of two parts. An EKF-based sparse V-SLAM system is first implemented. All landmark locations, the camera's current angle, and an as-yet-unspecified subset of previous camera poses are contained in the EKF's state vector. Researchers addressed the computational complexity issue with EKF SLAM by using so-called conditionally independent sub maps, which mostly have a constant run time complexity. They construct a dense point cloud from stereo in the second section. This is a local map, a dense point cloud. It is obtained through the process of filtering disparity values for individual image pixels. These disparities are densely reconstructed and expressed in local coordinate systems. Each past camera location tracked by the sparse V-SLAM crosses one of these local coordinate systems. Drift is fixed and previous camera poses are updated at loop closure. As a result, the local dense maps' global position is also updated. In this way, researchers propose a method to compute a dense local map that is embedded in the sparse SLAM algorithm. The dense local map is constantly updated, and the 3D global map is obtained after a closed loop. The experiments also show that dense maps may be more accurate than original stereoscopic reconstruction [6].

Some studies have shown that robotic odometers can greatly help autonomous robot movement, which can also help in the autonomous driving of cars. The feasibility of car autonomous driving based on SLAM technology can be proved through car-like robot experiments.

When a robot uses wheel encoders or other motion sensors to estimate its motion over time, the result of motion integration is referred to as odometry in robotics. Numerous odometry models have been developed, based on different kinds of sensors. One such model is Visual Odometry (VO), or odometry with cameras. A significant amount of uncertainty in robot navigation can be removed by improving the odometry model's accuracy, which also improves the mapping procedure. Primarily, mapping is significant in three areas:

- In order to facilitate course planning, obstacle avoidance, etc., maps are required.
- Many applications in mobile robotics aim to produce the map themselves.
- The correctness of the mapping has a major impact on the robustness and localization accuracy.

Among the most crucial components of mapping, loop closure enables the robot to identify a visited location and, consequently, optimize its estimated pose. The drifts are significantly reduced by closing the loop, allowing the robot to correct its odometry errors. The mapping process is typically included in the more recent visual odometry techniques, but the map may not be useful for path planning or other local task execution tasks. Thus, due to global map optimization, specifically the loop closure, SLAM varies from current odometry models (Cadena et al., 2016b).

Figure 1 show the practical application of SLAM technology for autonomous automotive driving.

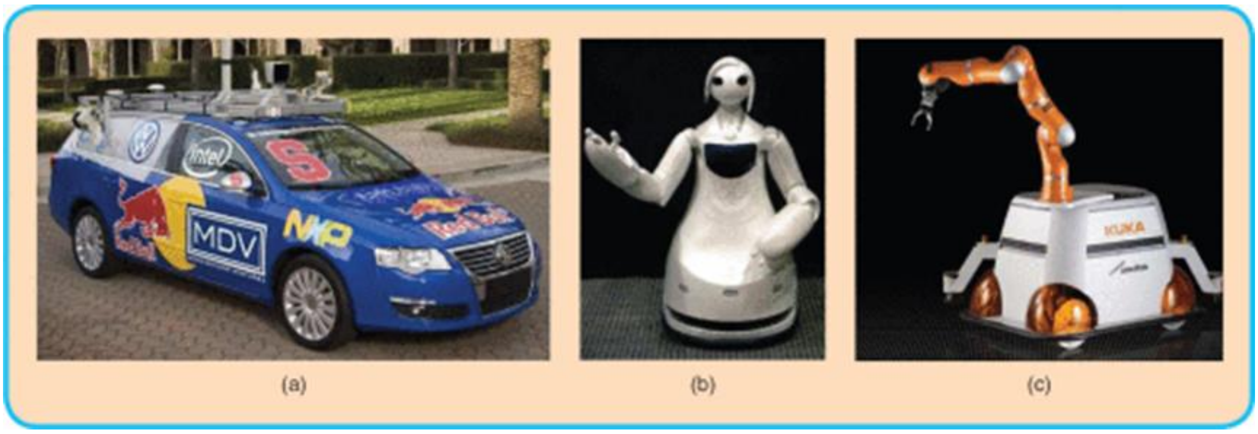

Fig. 1 A Tutorial on Graph-Based SLAM [1].

4. The drawbacks and solutions to autonomous driving based on SLAM technology

However, there are still drawbacks to autonomous driving based on SLAM technology. Location and hardware uncertainty are the two main categories of SLAM uncertainty that impact the car's capabilities and performance (Pirahansiah et al., 2013) [7]. Hardware errors and noises in the car's components lead to hardware uncertainty, which in turn leads to the extraction of incorrect and inaccurate information, which in turn leads to the inaccurate analysis of the pose, landmarks, and other calculating factors. The location of the moving car and the existence of multiple paths in environments lead to location uncertainty.

At the same time, due to defective sensors, which are a part of the robot's hardware, the sensory data somewhat deviates from reality. For example, the output of sensors can be affected by climate variations and the physical aspects of the surroundings. For instance, when utilizing cameras or other optical sensors, light shifts can impact perception. Bordoni and Damico thoroughly investigated noise in sensors (1990). Additionally, the calculated measurements are produced by various measurement models, which are mathematical relations that compute the relative information—which may not be accurate—between the states of the landmark and the sensory signals. Over time, the little perception errors accumulate, leading to system failure in long-term navigation. It is exacerbated by low-quality sensors. Therefore, observation error is a significant problem in SLAM that needs effective fixes [5].

Through M. W. M. G. Dissanayake's article, it is easy to conclude that the SLAM problem is unsolvable. As a result, the general SLAM problem has an answer, and it is possible to compute car position estimations simultaneously and produce a totally accurate map without any prior knowledge of vehicle or landmark locations. Beginning with Smith, Self, and Cheeseman's initial estimation-theoretic work, this research verifies the following three conclusions: As observations are made one after the other, any submatrix of the map covariance matrix has a monotonically falling determinant.; The landmark estimates become fully connected in the limit as the number of observations rises; Only the starting covariance in the vehicle location estimate determines, in the limit, the covariance associated with any given landmark location estimate.

These three results fully explain the map's steady state behavior and convergence properties. The map of relative locations is known with absolute precision as the vehicle moves through the environment, and the total uncertainty of the estimates of landmark locations decreases monotonically. In the limit, errors in any pair of landmark estimations exhibit full correlation. This implies that any other landmark on the map can be located with 100% accuracy given the precise location of any other landmark. The error in the absolute position estimate of each landmark (and consequently the entire map) approaches a lower bound that is solely controlled by the error present at the time of the initial observation as the map converges in the way that is explained in the text. At the same time, it is not difficult to see from the content displayed that solutions to common SLAM problems do exist.

As for the implementation of the simultaneous localization and map building algorithms, the researchers used the estimated landmark location covariance to perform the calculation. In order to solve the SLAM problem, the entire map covariance matrix must propagate. The cross-correlations in this map covariance matrix support the demonstrated convergence properties, which preserve knowledge of the relative relationships between landmark location estimates. If these cross correlations are not included, the SLAM problem loses its entire structure and the map building problem has inconsistent and divergent solutions [8].

Since local maps are being reviewed in hierarchical SLAM, it is also possible to enhance their estimations by applying the conventional EKF approach. It is well known that linearization errors cause this to yield a suboptimal result. Errors in linearization have little impact because the local maps are small. All other local map-based techniques are equivalent to hierarchical SLAM at this local level in that local maps and their relative transformations will converge to the actual solution. The primary distinction manifests itself when a loop is identified on a global scale. Similar to other two-level mapping systems, hierarchical SLAM also experiences suboptimality in its solution due to the division of the local and global levels. In order to preserve the independence between local maps, corrections made at the global level due to loop closure do not propagate back to the local level. These adjustments, though, are quite minor and have no bearing on the accuracy of local maps. It is possible to consistently determine a local map's absolute location, but the correlation between the local and global levels will be disregarded if this information only includes the locations of the local map's elements. That said, this discrepancy is probably not very significant. If necessary, this inconsistency can be avoided by recalculating the loop constraints and implementing this functionality at the global level. At the boundary between local maps, there is another reason of suboptimality. The consistency of a feature's absolute location is not required if it is seen in two adjacent maps. The local maps and the connection between them would become more precise if this restriction were applied, but the maps will start to correlate. Bailey speculates that disregarding these connections still yields consistent results, but gives no proof [9].

In hierarchical SLAM, the loop closing is enforced by a technique that maximizes the utilization of global information. The efficacy of the conducted experiments described in C. Estrada's paper appears to be dependent on the scan matching technique's accuracy, which yields accumulated heading errors of less than 1° [10].

5. Conclusion

In this paper, SLAM technology is briefly introduced. The experimental results of some researchers are summarized. The application of SLAM and the points that need to be improved are expounded. It is proved that SLAM system can indeed calculate dense maps more accurately than simply reconstructing parallax images. Many studies have shown that the system works reliably in a variety of scenarios, and SLAM can build a good map regardless of uncertainties such as camera shake. This is especially helpful in scenarios where the exogenous sensor is unreliable, such as when GPS is operating in a cluttered environment. These results are very useful for the application of SLAM to car autonomous driving.

Even though there are shortcomings in applying this technology, as mentioned in the article, there are ways to solve its drawbacks. Developing more advanced local dense mapping algorithms is the direction of future research, and if dense optical flow and parallax are integrated, dense flow algorithms can become more and more feasible. It is believed that in the future SLAM technology can be better and more convenient to be applied to the field of automatic driving, so that the car will be more accurate in real-time conditions such as road conditions, which makes automatic driving optimized, so that people can get rid of tired driving to the greatest extent.

References

[1] Grisetti G., Kümmerle R., Stachniss C., et al. A Tutorial on Graph-Based SLAM. IEEE Intelligent Transportation Systems Magazine, 2010, 2(4): 31-43.

[2] Taketomi T., Uchiyama H., Ikeda S. Visual SLAM algorithms: a survey from 2010 to 2016. IPSJ Transactions on Computer Vision and Applications, 2017, 9: 16.

[3] Lu F., Milios E. Globally consistent range scan alignment for environment mapping. Autonomous Robots, 1997, 4: 333-349.

[4] Li J, Bi Y, Lan M, et al. Real-time simultaneous localization and mapping for uav: A survey[C]//Proc. of International micro air vehicle competition and conference. 2016, 2016: 237.

[5] Taheri H., Xia Z. C. SLAM; definition and evolution. Engineering Applications of Artificial Intelligence, 2021, 97: 104032

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

[7] Pirahansiah F, Abdullah S N H S, Sahran S. Simultaneous localization and mapping trends and humanoid robot linkages[J]. Asia-Pacific Journal of Information Technology and Multimedia, 2013, 2(2): 27-38.

[8] Dissanayake M. W. M. G., Newman P., Clark S., et al. A solution to the simultaneous localization and map building (SLAM) problem. IEEE Transactions on Robotics and Automation, 2001, 17(3): 229-241

[9] Bailey T. Mobile robot localization and mapping in extensive outdoor environments. 2002.

[10] Estrada C., Neira J., Tardos J. D. Hierarchical SLAM: real-time accurate mapping of large environments. IEEE Transactions on Robotics, 2005, 21(4): 588-596