Off-road RGB-D SLAM and Path planning with GPS

Jueun Mun
Kyung Hee University
cindy4741@khu.ac.kr

Seongil Heo
Hankuk University of Foreign Studies
tjddlf101@hufs.ac.kr

Jiwoong Choi
Kyung Hee University
jwtiger22@khu.ac.kr

Jiwon Park
Kyung Hee University
overflow21@khu.ac.kr

Jaea Cho
Purdue University
cho418@purdue.edu

***Abstract*—This paper proposes robust RGB-D SLAM system and path planning algorithm for autonomous vehicles at the off-road environment. The consideration of limitations at the off-road was critical in this research. Adequate sensors (RGB-D camera, GPS) were choosed and SLAM methods were decided out of primary visual SLAM methods. At the path planning the priority of the path was determined in the light of the characteristics of off-road : slope, trunk, and rocks. GPS sensor conducts the primary role both at the SLAM system and path planning. GPS sensor compliments every three multithread of RGB-D SLAM : camera tracking, local mapping, and loop closing. The priorities for the optimal path were marked by GPS sensor to assign more weight. The process of making prototype and testing the system is being progressed.**

***Index Terms*—Autonomous vehicles, SLAM, path planning, off-road driving.**

I. Introduction

In recent years, Autonomous vehicles (AVs) are being developed as the time goes by and show promising performance. Modern automated vehicles have found application in numerous fields, including mobility, agriculture, and the space industry, where the capabilities of the driving situation have improved significantly. In Automated Driving Systems (ADS), there are various tasks linked and interacted with each other. Generally, ADS can be divided by three stages: (i) Perception stage, which understands the environment, such as location, tracking and pose; (ii) Decision stage, which generates an optimal action plan; and (iii) Control stage, which produces an actual motion. The purpose of each stage is achieved by several tasks. Simultaneous Localization and Mapping (SLAM), path planing and actuator are [10] in perception, decision, and control stage, respectively, and play significant roles in each of these phases. Therefore, several tasks show meaningful performance, especially in the on-road, which is paved road (covered with concrete) has lined on it. On the other hand, several methods for the on-road are not applicable to off-road since there are lack of the landmarks or lane markings, uncertainty as to which objects may be traversable or not, and variable terrain.

When using SLAM, it is required to choose and match specific points in each frame for localizing and mapping. The program chooses special points automatically based on their distinguishing features, such as corners and edges. In the off-road, especially farmland or forest, it is difficult to identify the specific points due to the similarity of the environment. Furthermore, slightly changes in the environment caused by natural features may disturb the process of creating local maps. The traditional method of path planning has some limitation in terms of adaptability to off-road situations. In most cases, the optimal path is selected based on the distance between the start point and the destination point. Nevertheless, in off-road terrain, the shortest path is not the optimal route, since the quality of each route varies, and some of them might not be appropriate for driving. Hence, it is imperative to choose an appropriate method for driving off-road.

Recently, there are many methods in SLAM and path planning to perform robustly in the off-road. The technique of SLAM using a camera with advantages such as low power consumption and high performance has become more prominent. However, using single camera cannot effectively adapt to weak dynamic off-road environment. To make the system robust in the complex environment, researchers use RGB-D cameras, which combine the depth and RGB information. In path planning, the sampling based planners are applied in the off-road environment to reduce the computational complexity such as BIT* [25] and ABIT* [25]. Still there are some limitation to use these methods in the off-road field.

It is well known that GPS data quality is affected by the landscape characteristics and the presence of buildings. The environment of the on-road usually has lots of buildings and may reduce the accuracy of the GPS data. On the other hand, the off-road environment is more possible to use the GPS sensor.

This paper presents a RGB-D SLAM that takes advantage of the GPS sensor to capture additional information about the environment and a path planning approach that takes the samples into consideration in order to find a safe and optimal path based on the characteristics of off-road vehicles. In contrast to pure RGB-D SLAM and path planning, these methods perform better in an off-road environment.

II. Related Works

A. Off-road environment

Off-road refers to roads that are not paved for cars such as farms, deserts, and mountains. Off-road has uneven terrain, strong texture, and irregular characteristics as there are no clearly structured landmarks in open areas without lanes, signals and rules [7]. Therefore there is a big difference between various types of environments such as mountains and deserts.

With this charaterstics, there are additional challenges to develop an autonomous driving system in an off-road environment: First of all, the class of semantic segmentation of data used in the off-road AV studies is not formalized because ther are a few similarities to various types of environments. A self-driving model developed in another environment cannot be directly transferred to another environment. AV studies conducted off-road require less data than those conducted on-road, however, have fewer data [8] available. Secondly, it should be able to respond to continuously changing even in the same environment. This is because obstacles are moved by the wind and plants continue to grow. Lastly, the condition of the ground should be checked. Accidents such as slipping or rollover of vehicles may occur due to ice or mud. Therefore, unlike on-road, off-road requires three-dimensional recognition of the environment, which increases the calculation time. Since it has to operate in real-time, the calculation time of the algorithm should be short [26].

B. RGB-D SLAM

RGB-D SLAM is a visual-based SLAM using not only 2D images but also depth information. Collecting accurate depth information is crucial, and it is done using depth sensor or depth estimation algorithm.

Alexandre Lopes et al ‘s survey [15] shows four types of sensor to collect depth information : Structured Light, Time of Flight (TOF), LiDAR, and Stereo camera sensing. They say at the outdoor environment, where there are light changes and the big variability of scene, LiDAR can be a good selection. At the indoor, structured light sensors are suitable.

The KinectFusion algorithm [20] is the first real-time SLAM algorithm using RGB-D sensor. It consists of four main steps : the measurement, pose estimation, reconstruction update, and suface prediction. Due to the lack of loop closing, this algorithm accumulated the errors.

SLAM++ is an object-oriented algorithm [22] There are four main steps: camera pose estimation, object insertion, and pose update, pose–graph optimization, and surface rendering. Even though this algorithm conducts loop detection and increased efficiency by object’s repeatability, it is suitable for already known scene, since it searches to identify objects in the current frame using the database information. [17]

Dense visual odometry is key-frame based SLAM [9]. Frame is made by certain interval and each frame are connected by graph. Each keyframe, which means a posture of a car becomes a vertex of a graph and edges are the transformation between keyframes.

Recently, the state of the art algorithms of visual SLAM consists multi threads : camera tracking, local mapping, and loop closing. PTAM [11] firstly divided visual SLAM into camera tracking and local mapping. Then, J. Engel et al, R. Mur-Artal et al added loop closing and global BA. The system of this research also consists of these threads.

Camera tracking extracts the map points contained in frames, and estimate the pose of the vehicle. Direct methods use the pixel intensity to match different frames. The mapped elements can be pixel maps. However, indirect methods extract features from each frame and use geometric constraints for matching. Feature descriptors often use intensity gradients to detect zones of interest [23]. Famous descriptors are Harris [5], SURF [1], SIFT [16], FAST [21], and ORB [18]. Considering the outdoor environment, where there are light changes, the variability and outlier of scene, it is recommended to choose direct method. DSO is a direct and sparse VO method and performs well at the environment.

Then, the pose of the vehicle is estimated by the map points. Comparing the map points between the current frame and the previous one, only meaningful frame that represents the certain location of the environment is decided as a keyframe.

Next, Local mapping thread builds a local map and optimizes the camera pose with the local map. This thread inserts new keyframes into a map, runs continuously, and extends the map.

Lastly, loop closing thread adjust the accumualted error. When the vehicle starts driving towards the unknown location, the error accumulates. As soon as the vehicle revisits the known location, system detect it and adjust the accumulated error. ORB-SLAM2 [19] builds a covisibility graph, and uses a subgraph of it for loop closing, which is made up of the spanning tree. In this process, the loop is detected by the bag-of-words library DBoW2 [3], which makes the long-term data association possible, and local Bundle Adjustment (local BA) is conducted, optimizing a local window of keyframes and points of the local map. After a loop closure, full BA is conducted and optimize all keyframes and points. The Levenberg– Marquardt method is used in this optimization process, implemented in g2o [13]. Besides, gradient descent method and Gauss-Newton method [27] is utilized for optimizing the cost function.

Meanwhile, RGB-D SLAM associate RGB-D data and inertial data for odometry adjustment. In this research, GPS data will be harnessed at the all multithreads to complement the inaccuracy at the outdoor.

C. Path planning

When the vehicles move the unknown environment, the strategy for trajectory is required. Path planning algorithm design the path so that the vehicles will be able to travel within the unknown environment with the least cost and without collision. In addition, when planning the movement of the vehicle, other information can be included, such as the movement of the obstacles or the movement of the other vehicles. Path planning algorithms can be classified into two types; Search based method and Sample based method.

1) Search based Method: Search based methods use the graph search method for finding the trajectory of the vehicle. In order to use the search based methods, the environment needs to be transformed into a discrete representation. The A* algorithm [6], which is similar with the Dijkastra's algorithm, guides users to the most optimistic states while saving significant amount of computation time compared to the Dijkstra's algorithm.The algorithm searches only those paths that succeed in connecting the nodes in the subgraph and selects the most efficient path based on the cost. Despite this, due to the discrete representation, A* only passes through the center of the gird. The accuracy depends on resolution of the grid [2]. Hybrid A* [2] provides a solution to this problem by ensuring that the path remains continuous state with each grid cell. This can make it possible to implement on the non-holonomic vehicle. Moreover, Hybrid A* predicts the motion of the vehicle depending on the various parameters of the vehicle such as the speed, steering angle and gear. D* algorithm [24] is proposed for finding the trajectory in the dynamic environment. D* generates the path considering the moving obstacles, which can easily re-plans using the cost due to the detected obstacles. However, D* has the high computational complexity, making it difficult to use in the large dynamic environment.

2) Sample based Method: For larger dynamic environments, sample-based methods are used by sampling the configuration without presenting a graph representation. While there are some sample-based methods capable of discretization, most methods search for the path in continuous space. RRT [14] creates a path in every region and expands to reach the desired destination. During this process, the path connects the nearest vertex selected according to the distance, and each vertex is generated at random. Also, the path expands more in the unexplored space, and this can make the path more optimal.The RRT Connect [12] combines two RRT method in which one starts from the starting point, and the other starts from the goal point, and they are connected by using heuristics. However, RRT Connect is not a asymptotically optimal planners due to the greedy tree connection [4]. BIT* [4] takes advantage of the search based methods by reducing the wasted computing effort from searching paths randomly in each process. BIT* searches for the path within each batch, and when the path is found, prunes the sample that are located outside of the path. In this way, the time complexity and the search space can be reduced. However, BIT* has some limitation to implement on the off-road vehicle since it only considers the distance, not the geographic feature of the road.

The proposed BIT* with GPS consider not only the distance of the path but also the geographic feature. In an off-road situation, our method facilitates the selection of an optimal path and a stable path for driving.

III. Methods

Acknowledgment

References

[1] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *European conference on computer vision*, pages 404–417. Springer, 2006.

[2] Dmitri Dolgov, Sebastian Thrun, Michael Montemerlo, and James Diebel. Practical search techniques in path planning for autonomous driving. *Ann Arbor*, 1001(48105):18–80, 2008.

[3] Dorian Gálvez-López and Juan D Tardos. Bags of binary words for fast place recognition in image sequences. *IEEE Transactions on Robotics*, 28(5):1188–1197, 2012.

[4] Jonathan D Gammell, Siddhartha S Srinivasa, and Timothy D Barfoot. Batch informed trees (bit*): Sampling-based optimal planning via the heuristically guided search of implicit random geometric graphs. In *2015 IEEE international conference on robotics and automation (ICRA)*, pages 3067–3074. IEEE, 2015.

[5] Chris Harris, Mike Stephens, et al. A combined corner and edge detector. In *Alvey vision conference*, volume 15, pages 10–5244. Citeseer, 1988.

[6] Peter E. Hart, Nils J. Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.

[7] Kaijin Ji, Huiyan Chen, Huijun Di, Jianwei Gong, Guangming Xiong, Jianyong Qi, and Tao Yi. Cpfg-slam: A robust simultaneous localization and mapping based on lidar in off-road environment. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 650–655. IEEE, 2018.

[8] Peng Jiang, Philip Osteen, Maggie Wigness, and Srikanth Saripalli. Rellis-3d dataset: Data, benchmarks and analysis. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1110–1116. IEEE, 2021.

[9] Christian Kerl, Jürgen Sturm, and Daniel Cremers. Dense visual slam for rgb-d cameras. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2100–2106. IEEE, 2013.

[10] Eungcheol Kim, J Kim, and Myoungho Sunwoo. Model predictive control strategy for smooth path tracking of autonomous vehicles with steering actuator dynamics. *International Journal of Automotive Technology*, 15(7):1155–1164, 2014.

[11] Georg Klein and David Murray. Parallel tracking and mapping for small ar workspaces. In *2007 6th IEEE and ACM international symposium on mixed and augmented reality*, pages 225–234. IEEE, 2007.

[12] James J Kuffner and Steven M LaValle. Rrt-connect: An efficient approach to single-query path planning. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*, volume 2, pages 995–1001. IEEE, 2000.

[13] Rainer Kümmerle, Giorgio Grisetti, Hauke Strasdat, Kurt Konolige, and Wolfram Burgard. g 2 o: A general framework for graph optimization. In *2011 IEEE International Conference on Robotics and Automation*, pages 3607–3613. IEEE, 2011.

[14] Steven M LaValle et al. Rapidly-exploring random trees: A new tool for path planning. 1998.

[15] Alexandre Lopes, Roberto Souza, and Helio Pedrini. A survey on rgb-d datasets. *arXiv preprint arXiv:2201.05761*, 2022.

[16] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[17] Andréa Macario Barros, Maugan Michel, Yoann Moline, Gwenolé Corre, and Frédérick Carrel. A comprehensive survey of visual slam algorithms. *Robotics*, 11(1):24, 2022.

[18] Raul Mur-Artal, Jose Maria Martinez Montiel, and Juan D Tardos. Orb-slam: a versatile and accurate monocular slam system. *IEEE transactions on robotics*, 31(5):1147–1163, 2015.

[19] Raul Mur-Artal and Juan D Tardós. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE transactions on robotics*, 33(5):1255–1262, 2017.

[20] Richard A Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J Davison, Pushmeet Kohi, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking. In *2011 10th IEEE international symposium on mixed and augmented reality*, pages 127–136. IEEE, 2011.

[21] Edward Rosten and Tom Drummond. Machine learning for high-speed corner detection. In *European conference on computer vision*, pages 430–443. Springer, 2006.

[22] Renato F Salas-Moreno, Richard A Newcombe, Hauke Strasdat, Paul HJ Kelly, and Andrew J Davison. Slam++: Simultaneous localisation and mapping at the level of objects. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1352–1359, 2013.

[23] Myriam Servières, Valérie Renaudin, Alexis Dupuis, and Nicolas Antigny. Visual and visual-inertial slam: State of the art, classification, and experimental benchmarking. *Journal of Sensors*, 2021, 2021.

[24] A. Stentz. Optimal and efficient path planning for partially-known environments. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, pages 3310–3317 vol.4, 1994.

[25] Marlin P Strub and Jonathan D Gammell. Advanced bit*(abit*): Sampling-based planning with advanced graph-search techniques. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 130–136. IEEE, 2020.

[26] Han Wang, Hongjun Zhang, Kun Wang, Chen Zhang, Chengxiang Yin, and Xingdang Kang. Off-road path planning based on improved ant colony algorithm. *Wireless Personal Communications*, 102(2):1705–1721, 2018.

[27] Robert WM Wedderburn. Quasi-likelihood functions, generalized linear models, and the gauss—newton method. *Biometrika*, 61(3):439–447, 1974.