SUGISAKA MASANORI

Journal of Robotics, Networking and Artificial Life
Vol. 10(2); September (2023), pp. 127–130
ISSN (Online): 2352-6386; ISSN (Print): 2405-9021
https://alife-robotics.org/jrnal.html

Research Article

Optical Sensors Fusion Approaches for Map Construction: A Review of Recent Studies

Alexander Eryomin, Ramil Safin, Tatyana Tsoy, Roman Lavrenov, Evgeni Magid
Department of Intelligent Robotics, Kazan Federal University, Kazan, Russia

ARTICLE INFO

Article History

Received 21 November 2022
Accepted 20 October 2023

Keywords

Data fusion
Mapping
SLAM
Machine learning

ABSTRACT

Map construction, or mapping, plays an important role in robotic applications. Mapping relies on inherently noisy sensor measurements to construct an accurate representation of a surrounding environment. Generally, individual sensors suffer from performance degradation issues under certain conditions in the environment. Sensor fusion enables to obtain statistically more accurate perception and to cope with performance degradation issues by combining data from multiple sensors of different modalities. This paper describes the latest developments in data fusion and state-of-the-art mapping methods using data fusion.

1. Introduction

Mapping is a process of constructing a map of an environment using robot perception. There exist multiple map representations, such as sparse point clouds, topological maps, and dense voxel grids. Mapping could be difficult due to adverse conditions (e.g., low lighting) and a presence of dynamic objects.

In practice, one sensor may not produce an accurate map due to inherent sensor issues. For example, data obtained from a Light Detection and Ranging sensor (LiDAR), contains errors that increase as the distance to an object increases. A beam hitting a poorly reflecting surface leads to its distortion and data loss, adverse weather conditions cause a performance decrease. On the other hand, radars are less vulnerable to environmental conditions, but have less accuracy due to large wave sizes - from centimeters to meters. Mono- and stereo cameras, in turn, depend on lighting conditions and require a precise intrinsic calibration to correct an image distortion. To obtain reliable maps, it is required to combine strengths of each sensor to cope with their weaknesses.

Sensor fusion, or data fusion, is a technique for combining data from multiple sensors in a way that enables to obtain a more reliable and accurate information about a system being measured.

An aggregation of measurements statistically improves an overall accuracy using multiple sources of information. For instance, one can combine range data from LiDAR, radar, sonar, and RGB-D cameras.

Fig. 1 shows one of sensor configurations mounted on an autonomous car for a data fusion. LiDAR and radar estimate a distance to objects, and a camera provides color information to form a point cloud. This approach is actively used in robotics, particularly, in Simultaneous Localization and Mapping (SLAM) [1].

2. Sensor Fusion

A high-level architecture of a sensor fusion system consists of the following components:

1. ***sensors*** that independently measure an observed quantity;
2. mathematical ***models*** that convert the observed quantity into a target value (e.g., calculating a position from distance measurements);

Corresponding author's E-mail: safin.ramil@it.kfu.ru

3. an ***inference algorithm*** that calculates a resulting target value by combining data from all sensors.

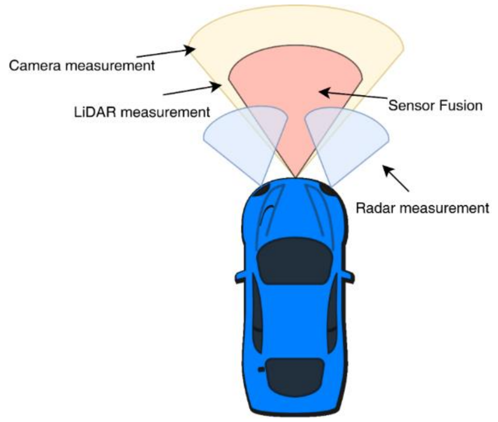

Fig. 1. A schematic representation of data fusion from a camera, LiDAR, and radar for a self-driving car.

Modern data fusion approaches can be classified according to their mathematical tools: traditional and based on machine learning (ML) [2]. Methods can also be divided according to a data fusion level [3].

1. a signal level fusion – integrating raw data from sensors through a direct relationship between measurements;
2. a feature level fusion – combining data from sensors using detected features;
3. a decision-making level fusion – an inference of a single solution from multiple hypotheses based on data from different sensors.

Signal-level fusion combines raw data. The most applicable data correlation technique performs an extrinsic calibration of a system consisting of several sensors. A calibration problem solution usually consists of finding a homogeneous transformation matrix, which is a transformation between the coordinate systems of sensors. Extrinsic calibration methods are divided into traditional (using a fiducial target) and self-calibration methods (using scene data). Traditional methods use physical objects as a marker that introduces geometric constraints. For example, it is common to use a planar board (chessboard, ArUco, Apriltag, etc.) for an extrinsic calibration of a LiDAR-camera system. In this case, a point-to-plane constraint is used, which states that points obtained from the LiDAR lie on a plane of a planar board. There are also solutions based on more complex calibration objects, such as a V-shaped calibration object, which provides six point-to-plane constraints [4]. The paper [5] presents a method of finding extrinsic parameters through scanning of spherical target. Self-calibration, or online calibration, is an approach that allows calibrating a sensor system based on the information about the environment.

Some of the approaches are based on ML. For instance, in [6], the authors present a neural network which refines calibration parameters. An image, a depth map, and an initial solution are issued to the RegNet. An output matrix is used to compute an optimal transformation matrix. Data from a LiDAR and a camera are combined using features that are calculated by passing through convolutional layers. This approach characterizes the feature level fusion, but since the output presents extrinsic parameters, we consider the method as the signal-level. In [7], a transformer is added to the convolutional layers.

Traditional approaches are also used in online calibration. The paper [8] presents a calibration using Structure-from-Motion (SfM) methods. SfM allows obtaining an initial guess for intrinsic and extrinsic parameters. Then the bundle adjustment method corrects a solution by minimizing a reprojection error.

Feature-level fusion combines detected features from different sensors. In ML features are primarily extracted from dense data (e.g., from image pixels) using convolutional neural networks (CNNs). In [9] a CNN is used as an encoder, obtained features are decoded to obtain normals and an initial dense depth of a scene. In [10] non-local neighbors and their affinities of each pixel are used instead of normals. A feature extraction enables one to semantically describe a map. In [11] a semantic scene completion network (SSCNet) is proposed, which simultaneously outputs an occupancy grid and semantic labels for all voxels. The paper [12] presents a multi-modal deep feature architecture of a neural network for a point cloud refinement from LiDAR using a single image.

Decision-level fusion combines decisions of multiple classifiers into a common decision about a map. For example, one can semantically combine objects detected from RGB-D and LiDAR.

3. Multi-modal Map Construction

Data fusion is used to construct maps based on data of different modalities. A mapping task is generally not independent and is solved along with a localization in SLAM. The paper [13] presents a SLAM algorithm for unmanned ground vehicles (UGV) based on an estimation of an odometry using data from 2D LiDAR. To refine the odometry and correctly display an environment, key frames obtained using LiDAR and a stereo camera are combined by finding cloud point correspondences. The correspondence between the clouds of related keyframes is calculated using the Iterative Nearest Point algorithm. The map is generated as a Probabilistic Occupancy Grid (POG) (Fig. 2). This representation of a map is considered as a basic Bayesian integration method [14]. The occupancy grid map divides an environment into cells of equal size. Each cell carries information about the probability of its occupancy. The Bayesian rule allows updating initially constructed maps. Computational performance is highly dependent on a map resolution and on a number of degrees of freedom. In practice, 2D maps are often considered due to a lower complexity and easier data fusion [15]. In [13] a 2D map is created, which simplifies calculations complexity.

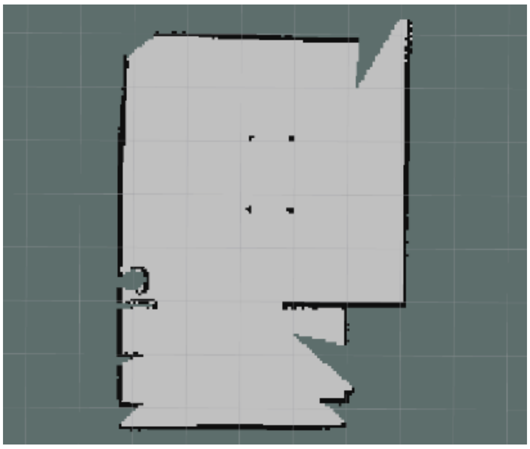

Fig. 2. The representation of occupancy grid.

In [16] a direct multi-sensor SLAM is used. To compare data from a rangefinder and a camera, the traditional approach of extrinsic calibration with a chessboard is utilized. Data loss can be avoided with a signal fusion approach. A reconstructed map is displayed as a 3D point cloud, where each point has cartesian coordinates and a color. Further, point clouds can be transformed by a surface reconstruction into polygonal or triangle mesh models [17], non-uniform rational B-spline surface models [18] or CAD models [19]. To construct and refine a map, state estimation techniques are employed [20], such as recursive filtering approaches. In [21] it is proposed to use Unscented Particle Filter and Unscented Bingham Filter to reduce an uncertainty.

In [22] authors use data from a LiDAR, an inertial measurement unit (IMU) and encoders for localization and mapping of a UGV. Measurements from the IMU and encoders are combined to assess a robot's pose. Range data is used to estimate odometry. In [23] a robot's pose is estimated using data from an IMU and a stereo camera. Received information is associated with rangefinder data and a map is constructed. In [24] GPS data is used for a localization in addition to an IMU and a rangefinder. Sensor data is issued to the Kalman filter – a position, a velocity and an attitude are determined.

4. Conclusion

In conclusion, most of the reviewed sensor fusion mapping techniques are based on traditional methods, such as Extended Kalman Filtering and particle filters. The feature-level fusion is a more time-efficient approach, which involves feature extraction and matching, which improves computational performance. In the SLAM systems case, bundle adjustment optimization methods prevail. Most of the time, laser scanners are used to improve camera depth estimation capabilities. Machine learning solutions allow learning depth from images but lack accuracy and reliability.

Acknowledgements

The reported study was funded by the Russian Science Foundation (RSF) and the Cabinet of Ministers of the Republic of Tatarstan, research project No.22-21-20033.

References

1. Xu X. et al. A review of multi-sensor fusion slam systems based on 3D LIDAR // Remote Sensing. MDPI, 2022. Vol. 14, № 12. P. 2835.
2. Velasco-Hernandez G., Barry J., Walsh J. Autonomous driving architectures, perception and data fusion: A review // International Conference on Intelligent Computer Communication and Processing. IEEE, 2020. P. 315–321.
3. Cocchi M. Data fusion methodology and applications. Elsevier, 2019.
4. Dong W., Isler V. A novel method for the extrinsic calibration of a 2D laser rangefinder and a camera // IEEE Sensors Journal. IEEE, 2018. Vol. 18, № 10. P. 4200–4211.
5. Tóth T., Pusztai Z., Hajder L. Automatic LiDAR-camera calibration of extrinsic parameters using a spherical target // 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020. P. 8580–8586.
6. Schneider N. et al. RegNet: Multimodal sensor registration using deep neural networks // 2017 IEEE intelligent vehicles symposium (IV). IEEE, 2017. P. 1803–1810.
7. Yuan K., Guo Z., Wang Z.J. RGGNet: Tolerance aware LiDAR-camera online calibration with geometric deep learning and generative model // IEEE Robotics and Automation Letters. IEEE, 2020. Vol. 5, № 4. P. 6956–6963.
8. Tu D. et al. Multi-Camera-LiDAR Auto-Calibration by Joint Structure-from-Motion // 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2022. P. 2242–2249.
9. Xu Y. et al. Depth completion from sparse lidar data with depth-normal constraints // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019. P. 2811–2820.
10. Park J. et al. Non-local spatial propagation network for depth completion // Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIII 16. Springer, 2020. P. 120–136.
11. Song S. et al. Semantic scene completion from a single depth image // Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. P. 1746–1754.
12. Yu X. et al. MMDF: Multi-Modal Deep Feature Based Place Recognition of Mobile Robots With Applications on

Cross-Scene Navigation // IEEE Robotics and Automation Letters. IEEE, 2022. Vol. 7, № 3. P. 6742–6749.
13. Jin Z. et al. A multisensor data fusion approach for simultaneous localization and mapping // 2019 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE, 2019. P. 1317–1322.
14. Yue Y. et al. Hierarchical probabilistic fusion framework for matching and merging of 3-d occupancy maps // IEEE Sensors Journal. IEEE, 2018. Vol. 18, № 21. P. 8933–8949.
15. Gupta A., Fernando X. Simultaneous localization and mapping (slam) and data fusion in unmanned aerial vehicles: Recent advances and challenges // Drones. MDPI, 2022. Vol. 6, № 4. P. 85.
16. Shin Y.-S., Park Y.S., Kim A. Direct visual slam using sparse depth for camera-lidar system // 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018. P. 5144–5151.
17. Ran H., Liu J., Wang C. Surface representation for point clouds // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022. P. 18942–18952.
18. Sharma G. et al. Parsenet: A parametric surface fitting network for 3d point clouds // Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VII 16. Springer, 2020. P. 261–276.
19. Gothandaraman R., Jha R., Muthuswamy S. Reflectional and rotational symmetry detection of cad models based on point cloud processing // 2020 IEEE 4th Conference on Information & Communication Technology. IEEE, 2020. P. 1–5.
20. Terblanche J., Claassens S., Fourie D. Multimodal navigation-affordance matching for slam // IEEE Robotics and Automation Letters. IEEE, 2021. Vol. 6, № 4. P. 7728–7735.
21. Pessanha S. N., Lobo V., Bernardino A. Unscented Particle Filters with Refinement Steps for UAV Pose Tracking // Journal of Intelligent & Robotic Systems. Springer, 2021. Vol. 102, № 2. P. 52.
22. Su Y. et al. GR-LOAM: LiDAR-based sensor fusion SLAM for ground robots on complex terrain // Robotics and Autonomous Systems. Elsevier, 2021. Vol. 140. P. 103759.
23. Jia Y. et al. Lvio-fusion: A self-adaptive multi-sensor fusion slam framework using actor-critic method // 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021. P. 286–293.
24. Baharom A.K. et al. Towards modelling autonomous mobile robot localization by using sensor fusion algorithms // 2020 IEEE 10th international conference on system engineering and technology (ICSET). IEEE, 2020. P. 185–190.

Authors Introduction

Mr. Alexander Eryomin

He received a BSc degree from Siberian Federal University in 2022. Currently, he is a first-year student of Master degree program in Intelligent Robotics at the Institute of Information Technology and Intelligent Systems at the Kazan Federal University.

Mr. Ramil Safin

He is a Research Assistant and Ph.D. student in Kazan Federal University. His research interests are computer vision, mobile robotics, and SLAM.

Ms. Tatyana Tsoy

She graduated from the University of Tsukuba in 2012. Since 2018 she has been a PhD student in Robotics at the Institute of Information Technology and Intelligent Systems of Kazan Federal University.

Assistant Professor Roman Lavrenov

He received his PhD from Kazan Federal University in 2020. Currently, he works as an Assistant Professor at Kazan Federal University, Russia. His research interests are path planning, ROS/Gazebo, and mobile robotics.

Professor Evgeni Magid

He is A Professor, a Head of Intelligent Robotics Department and a Head of Laboratory of Intelligent Robotic Systems at Kazan Federal University, Russia. Professor at HSE University, Russia. Senior IEEE member. He authors over 200 publications.