**SegDF: Segment Based Dynamic Filter in 3D Point Cloud**

*Abstract*

Dynamic objects in lidar data often leave unwanted traces in 3D point clouds, affecting the quality of maps and localization performance. We propose a novel static reconstruction method, called SegDF, for focusing on filtering the dynamic objects on the ground with high mobility in urban environment. Given pairs of registered scans, we build a curved-voxel map assisted by sensor intensity to cluster them into instances and obtain semantics through geometric verification. Meanwhile, we represent the occupancy state of the volume of space by curved-voxels to remove dynamic points in object level. Furthermore, our approach tightly couples the instance segmentation and dynamic detection to refine the imperfect clustering and improve the removal accuracy in the process of building a static instance map. We validate SegDF on the KITTI dataset using SemanticKITTI as ground truth and prove that it works well in real urban streets.

*Ⅰ. INTRODUCTION*

Recent advances in 3D light detection and ranging (LiDAR) mapping have been reported leveraging LiDAR odometry [1, 2, 3], place recognition [4, 5, 6], and simultaneous localization and mapping (SLAM) [7, 8, 9]. Most autonomous robotic systems always assume that all observations are available, but only relying on static objects while rejecting dynamic objects as outliers can avoid confusion and achieve robustness.

During a process of SLAM using lidar in urban sites containing various dynamic objects like vehicles, pedestrians and so forth[14], thus we encounter non-static points with low interpretability in a raw scan data. The changes of maps are equivalent to the diversity of point cloud between two registered scans, and existing algorithms [10, 11, 12, 13] also utilize this idea to remove dynamic objects. Unfortunately, the difference of maps are divided into two categories: a) moving object and b) residual caused by restricted view as in Fig.1(体现遮挡和动态残影), the former can be reflected as ghost trail effect [10, 13] by sequential accumulations of the scan data, but the latter might be misjudged as false dynamic points by the association errors in local maps. To tackle this problem, we propose a segment based dynamic filter, which processes a batch of measurements from one single-session SLAM and, unlike sequential Bayesian updates of [15], focuses on the post-processing role, because our main purpose is to construct a certain length of static object map without concerning the processing speed, and to ensure that actual dynamic objects are strongly erased even if some false negatives occur (i.e., points on actual static objects are wrongly erased).

In this paper, our method divides the process into two stages: a) instance segmentation assisted by sensor intensity and b) dynamic filtering in object level. It is worth mentioning that all scan data used in the post-processing process are generated from Lio-sam [3], and we consider the process errors due to imperfect pose estimation is tolerable. Firstly, we fill the pre-processed original point cloud into the coordinate system built by curved-voxel and perform clustering segmentation, which is similar to [16], but the difference is that we utilize sensor intensity calibrated additionally to refine the result and classify instances through geometric verification. The assistance of intensity makes up for the deficiency of clustering in [16], and cluster the distant divisive point clouds from the same object caused by the occlusion problem. Secondly, according to the semantics, we selectively execute dynamic detection on the objects with high mobility (e.g., moving objects contacted to the ground and unknown instance after geometric distinguishing). Furthermore, the two mechanisms are coupled tightly in the process of object tracking to refine the imperfect clustering and improve the tracking and removal accuracy. Therefore, the multi-frame data is integrated to compensate for the insufficient observation of objects in different screening.

Above all, we propose a novel post-processing reconstruction method, a segment based dynamic filter in 3D point cloud, called SegDF. Our contributions are threefold:

• An instance segmentation method based on curved-voxel clustering assisted by sensor intensity calibrated and geometric verification. (§III-B, §III-C)

• A dynamic filter in object level based on curved-voxel occupancy searching and instance tracking. (§III-D)

• A novel initialization method for instance map and a tight coupling scheme for instance refine and dynamic removal. (§III-E)

Ⅱ. RELATED WORK

A. Panoptic segmentation

As reported in [17], lidar panoptic segmentation is an ensemble of both the semantic segmentation for static stuff and the instance segmentation for countable objects. Though many researches [18, 19] directly dive into the deep learning solutions, despite the semantic classification part, point cloud clustering is a long-existing research topic that also has a chance to contribute as part of the panoptic task. Furthermore, geometric features [20, 21] can be used to distinguish clusters to obtain a low-level semantics.

Generally, the clustering methods of point clouds can be divided into three categories: a) clustering with Euclidean distance, b) clustering with supervoxels or superpoints and c) clustering on range image.

**Clustering with Euclidean Distance**. Using the Euclidean distance to cluster points is a straightforward idea. A novel ground segmenting algorithm was proposed in [22], other non-ground points were clustered with voxelized Euclidean neighbors. Following this philosophy, most clustering algorithms often extract the ground in advance to reduce calculation. In [23], researchers provided a probabilistic framework to incorporate not only the Euclidean spatial information but also the temporal information from consecutive frames. It firstly takes account of the shortcomings of European clustering and fuses the continuous frame information to improve the result, but it is still a point-to-point search, which makes it redundant.

**Clustering with Supervoxels or Superpoints**. Inspired by the concept of superpixels from the traditional image processing [24], some researchers are interested in finding super voxels or super points [25, 26] in the Euclidean space. Unlike superpixels using color or texture due to the sensor type, point cloud is hard to give these manual units more valuable information getting rid of Euclidean distance. In [16], the algorithm attains fine discriminations by considering three important aspects for clustering 3D LiDAR points: distance from the sensor, directional resolutions, and rarity of points to deal with sparse 3D point clouds. This method is similar to ours, but it still relies solely on geometric information and cannot successfully cluster point clouds that are split due to occlusion.

**Clustering on Range Image**. To improve the rate of clustering, many researchers explored more clues aiming at finding better criteria to separate neighbor points belong to different clusters. Representatively, in [27, 28], the angle formed by two adjacent laser beams is considered to construct the discriminator. To make the algorithm fast enough for real-time applications, authors of [27] worked on the 2D range image representation of the LiDAR point cloud. However, it is still a depth-dominated clustering method, especially when affected by the angle of view or the scanning line is sparse, it cannot segment the close objects or cluster the distant objects well.

B. Dynamic object removal

Reference

|  |  |
| --- | --- |
| [1] | Zhang, Ji, and Sanjiv Singh. "LOAM: Lidar odometry and mapping in real-time." Robotics: Science and Systems. Vol. 2. No. 9. 2014. |
| [2] | Shan, Tixiao, and Brendan Englot. "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain." 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018. |
| [3] | Shan, Tixiao, et al. "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping." 2020 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, 2020. |
| [4] | Kim, Giseop, and Ayoung Kim. "Scan context: Egocentric spatial descriptor for place recognition within 3d point cloud map." 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018. |
| [5] | Wang, Han, Chen Wang, and Lihua Xie. "Intensity scan context: Coding intensity and geometry relations for loop closure detection." 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020. |
| [6] | Dubé, Renaud, et al. "Segmatch: Segment based place recognition in 3d point clouds." 2017 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2017. |
| [7] | Durrant-Whyte, Hugh, and Tim Bailey. "Simultaneous localization and mapping: part I." IEEE robotics & automation magazine 13.2 (2006): 99-110. |
| [8] | Bailey, Tim, and Hugh Durrant-Whyte. "Simultaneous localization and mapping (SLAM): Part II." IEEE robotics & automation magazine 13.3 (2006): 108-117. |
| [9] | Cadena, Cesar, et al. "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age." IEEE Transactions on robotics 32.6 (2016): 1309-1332. |
| [10] | Pagad, Shishir, et al. "Robust method for removing dynamic objects from point clouds." 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020. |
| [11] | Schauer, Johannes, and Andreas Nüchter. "The peopleremover—removing dynamic objects from 3-d point cloud data by traversing a voxel occupancy grid." IEEE robotics and automation letters 3.3 (2018): 1679-1686. |
| [12] | Kim, Giseop, and Ayoung Kim. "Remove, then revert: Static point cloud map construction using multiresolution range images." 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2020. |
| [13] | Lim, Hyungtae, Sungwon Hwang, and Hyun Myung. "ERASOR: Egocentric ratio of pseudo occupancy-based dynamic object removal for static 3D point cloud map building." IEEE Robotics and Automation Letters 6.2 (2021): 2272-2279. |
| [14] | Behley, Jens, et al. "Semantickitti: A dataset for semantic scene understanding of lidar sequences." Proceedings of the IEEE/CVF international conference on computer vision. 2019. |
| [15] | Pomerleau, François, et al. "Long-term 3D map maintenance in dynamic environments." 2014 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2014. |
| [16] | Park, Seungcheol, et al. "Curved-voxel clustering for accurate segmentation of 3D LiDAR point clouds with real-time performance." 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019. |
| [17] | Zhao, Yiming, Xiao Zhang, and Xinming Huang. "A technical survey and evaluation of traditional point cloud clustering methods for lidar panoptic segmentation." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021. |
| [18] | Gasperini, Stefano, et al. "Panoster: End-to-end panoptic segmentation of lidar point clouds." IEEE Robotics and Automation Letters 6.2 (2021): 3216-3223. |
| [19] | Hong, Fangzhou, et al. "Lidar-based panoptic segmentation via dynamic shifting network." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021. |
| [20] | Wohlkinger, Walter, and Markus Vincze. "Ensemble of shape functions for 3d object classification." 2011 IEEE international conference on robotics and biomimetics. IEEE, 2011. |
| [21] | Weinmann, Martin, Boris Jutzi, and Clément Mallet. "Semantic 3D scene interpretation: A framework combining optimal neighborhood size selection with relevant features." ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2.3 (2014): 181. |
| [22] | Douillard, Bertrand, et al. "On the segmentation of 3D LIDAR point clouds." 2011 IEEE International Conference on Robotics and Automation. IEEE, 2011. |
| [23] | Held, David, et al. "A Probabilistic Framework for Real-time 3D Segmentation using Spatial, Temporal, and Semantic Cues." Robotics: Science and Systems. Vol. 12. 2016. |
| [24] | Achanta, Radhakrishna, et al. "SLIC superpixels compared to state-of-the-art superpixel methods." IEEE transactions on pattern analysis and machine intelligence 34.11 (2012): 2274-2282. |
| [25] | Ben-Shabat, Yizhak, et al. "Graph based over-segmentation methods for 3d point clouds." Computer Vision and Image Understanding 174 (2018): 12-23. |
| [26] | Landrieu, Loic, and Mohamed Boussaha. "Point cloud oversegmentation with graph-structured deep metric learning." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019. |
| [27] | Bogoslavskyi, Igor, and Cyrill Stachniss. "Fast range image-based segmentation of sparse 3D laser scans for online operation." 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2016. |
| [28] | Yuan, Xia, Yangyukun Mao, and Chunxia Zhao. "Unsupervised segmentation of urban 3d point cloud based on lidar-image." 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO). IEEE, 2019. |
| [29] |  |
| [30] |  |
| [31] |  |
| [32] |  |
| [33] |  |