Semantic Grid Map based LiDAR Localization in Highly Dynamic Urban Scenarios

Chenxi Yang[1], Lei He[1], Hanyang Zhuang[2], Chunxiang Wang[1], Ming Yang[12]

***Abstract*— Change-over-time objects such as pedestrians and vehicles remain challenging for scan-to-map pose estimation using 3D LiDAR in the field of autonomous driving because they lead to incorrect data association and structural occlusion. This paper proposes a novel semantic grid map (SGM) and corresponding algorithms to estimate the pose of observed scans in such scenarios to improve robustness and accuracy. The algorithms consist of a Gaussian mixture model (GMM) to initialize the pose, and a grid probability model to keep estimating the pose in real-time. We evaluate our algorithm thoroughly in two scenarios. The first scenario is an express road with heavy traffic to prove the performance towards dynamic interferences. The second scenario is a factory to confirm the compatibility. Experimental results show that the proposed method achieves higher accuracy and smoothness than mainstream methods, and is compatible with static environments.**

I. Introduction

3D-LiDAR-based pose estimation is one of the most widely used on-line vehicle self-localization methods in the Global Navigation Satellite System (GNSS) signal denied or disturbed environments for autonomous driving. Data association from the observed scans to the pre-defined environmental map (scan-to-map) is the most critical step for such approaches, and the association reliability mainly determines the system performance. Dynamic interference in the observed scans has been a long-term challenge for data association in two aspects: (1) It provides time-varying features not present in the map that lead to incorrect data association; (2) It occludes extensive environmental features that lead to data association quantity reduction.

In order to suppress such impacts, one approach is to perceive the presence of incorrect data association during the pose convergence iteration and thus eliminate dynamic objects. However, this approach relies on the fact that the correct data association takes the dominant effect so that the outliers can be identified based on their significant distance error. Therefore, it's time-consuming and relatively unreliable. In this study, we focus on another approach, which is to directly exclude dynamic objects from the observed scans before data association. Our idea is to ensure the stationary status of the data association candidates by introducing semantic features. Such features should be widespread and generally static in urban environments.

This work is supported by the National Natural Science Foundation of China (U1764264/61873165).

[1]Department of Automation, Shanghai Jiao Tong University, Shanghai, 200240; Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, 200240 mingyang@sjtu.edu.cn.

[2]University of Michigan - Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai, 200240, China.

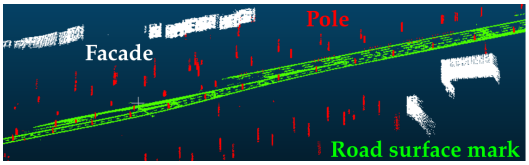


Fig. 1. The semantic point cloud map (SPCM) used in this study contains a sparse point cloud with only three kinds of static semantic features, which are poles, facades, and road surface marks.

Yu et al. [1] propose a semantic alignment method for city-scale LiDAR data by the data association limited to six kinds of semantic objects extracted from two dense point cloud maps. The features include facades, roads, poles, cars, segments, and lines. Their method achieves higher alignment accuracy than the mainstream methods. However, these semantic features (i.e., cars) are not designed to solve the challenges of dynamic scenarios, and a single scan is too sparse to extract segments and lines. Their map-to-map alignment method also faces a considerable gap to meet the real-time requirement of the vehicle self-localization. Inspired by their idea, we extract the static semantic features of poles, facades, and road surface marks using a dense semantic segmentation method [2]. Fig. 1 shows the semantic point cloud maps (SPCM) generated in this process.

Comparing to the inter-frame LiDAR-odometry, it is more accurate and robust to utilize the semantic information in real-time localization based on pre-defined maps with global consistency. However, there are several unique challenges remaining for the scan-to-map localization of this study: (1) The relative structural loss of the observed scan comparing with the pre-defined map is much more significant than that in inter-frame scans of the LiDAR-odometry; (2) Computing efficiency is highly required to reach real-time performance; (3) The pose initialization needs to be fast enough to complete initial localization in large scale maps.

In this study, we propose a novel semantic grid map (SGM) based on the SPCM, in order to improve the scan-to-map localization by alleviating the aforementioned three challenges. The semantic categories and the corresponding probabilities are assigned to each grid. By projecting the SPCM into SGM, we significantly speed up the calculation while guaranteeing the robustness towards dynamic interference. We realize a Gaussian mixture model (GMM) to initialize the pose of the observed scan in the SGM. After the initialization, we design a grid probability model to keep track of the vehicle in the SGM. We evaluate our method on an express road with heavy traffic. In both the pose

initialization and real-time localization, the proposed SGM and corresponding algorithms outperform the mainstream methods in precision and calculation speed. We also apply our method in a factory with generally static environment to confirm the compatibility.

The rest of the paper is organized as follows. Section II reviews related work. Section III introduces the SGM representation. Section IV describes our localization algorithms in detail. Section V provides the experimental evaluation of our method. Finally, Section VI concludes this paper.

II. Related Work

LiDAR localization approaches based on pre-defined maps can be classified according to their data association methods.

Point-based methods represented by Iterative Closest Points (ICP) [3] directly associate the observed points to the points on the map and get the result by converging the point-to-point optimization functions. Normal Distributions Transform (NDT) [4] transforms localization into a probability problem to solve and increases the robustness. Deschaud [5] proposed the IMLS-SLAM method to solve the data association problem by least-squares optimization and achieved high accuracy on the KITTI database. However, the point-based methods generally have poor real-time performances, therefore not suitable for on-board applications.

Gird-based methods can improve real-time performance by shrinking the map size. Levinson et al. [6] achieved a centimeter-level localization accuracy based on the probabilistic grid whereby every cell is represented as its own Gaussian distribution over remittance values. Wan et al. [7] proposed a similar method that is widely used in autonomous driving projects. Yang et al. [8] relieved ICP's local minima problem by combining a branch-and-bound (BnB) scheme. These methods can achieve high localization accuracy with great calculation speed in low dynamic scenes. However, just like the point-based methods, once the static correct data associations don't take the dominant effect, such methods will also suffer a severe accuracy loss.

Feature-based methods are currently the mainstream approach. Such methods extract abstract geometric structures such as lines and planes [9]. Generalized ICP (GICP) [10] realizes a plane-to-plane strategy that adopts the covariance matrices of the local surfaces to match with the point cloud. The Normal ICP (NICP) [11] assigns local geometric information of normal and curvature to the points to enrich the matching dimensions to improve the robustness further. LiDAR Odometry and Mapping (LOAM) [12], as one of the state-of-the-art methods, extracts edges and planes with great accuracy from the relatively sparse point clouds. Shan et al. [13] proposed a lightweight and ground-optimized LOAM variant that improved both speed and accuracy. Although these methods emphasize the structure of the objects, dynamic objects such as vehicles can also form local structures with strong consistency. Therefore, although feature-based methods are more robust to dynamic interference, they still cannot overcome the challenges brought by high-dynamic scenarios.

Descriptor-based methods cluster the point cloud into blocks and calculate the similarity between the observed scans and the maps based on the geometric measurement criteria. Dubé et al. [14] trained a similarity criteria. Lu et al. [15] designed a deep learning network to learn the point cloud characteristics and established the corresponding descriptor. Both these two methods efficiently improved the global localization performance, however, they have relatively slow calculation speed and poor interpretability.

To the best of our knowledge, the semantic category is one of the few (if not the only) features that can directly exclude the dynamic objects from the data association. Compared with the use of semantic cues in image-to-map registration tasks[16], LiDAR-based scan-to-map registration is more challenging due to the sparse information. Pole-like objects [17], [18], [19] and road surface marks [20], [21], [22] are often used for data association as their semantics strongly indicate these objects are static. In specific scenarios, these methods can eliminate the drawbacks caused by dynamic objects. But, relying on a single semantic feature will often result in a localization failure because of the occlusion and lack of structure. As introduced in the last section, Yu et al. [1] proposed a semantic alignment method that combined multiple semantic features to achieve higher localization accuracy. Parkison et al. [23] proposed a localization method based on the high-precision semantic segmentation of the dense point cloud. Chen et al. [24] computed semantic segmentation results in point-wise labels for the whole scan, allowing them to build a semantically-enriched map with labeled surfels. The global semantic segmentation process in these methods is time-consuming even on high-performance processors, therefore, it is almost impossible for on-line real-time applications.

III. Semantic Grid Map Representation

To accurately estimate the vehicle position, sufficient pose constraints from various directions and elevations are necessary. However, due to the sparseness of the point cloud, the static semantics extractable from a single LiDAR scan is relatively limited. For high-layer and ground semantics, facades and road surface marks are two robust static ones widespread in urban scenarios. However, in the middle layer, where the dynamic interferences are the most severe, it's typically challenging to find such features. Our idea is to strictly limit the static semantics, so as to distinguish them from the potentially dynamic ones effectively. Therefore, we choose only pole-like features, which implies that the objects are tree trunks or telephone poles.

Fig. 2 demonstrates the data structure of the proposed semantic grid map. Each grid is represented by the category determined by the semantic feature that has the most points, and the corresponding probability that is the proportion within the total points of this grid. Since the wrong data associations often occur at the boundaries of different categories (such as poles at the edges of the facade), the introduction of probability can weaken such impact. In some rare cases, SPCM contains some invalid semantic points, such as poles

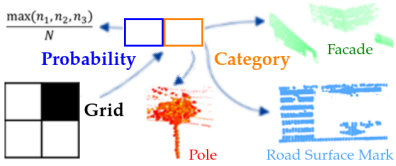


Fig. 2. Semantic grid map representation. Each occupied grid contains the information of the category and the corresponding probability.

formed by reeds. Although such points generally belong to static objects, it's obvious that they are not stable in the long term. What these features have in common is that their point clouds are more sparse than those intrinsically static features. Therefore, this paper removes such structures by a point number threshold.

IV. Localization

In this section, we describe our algorithm for the on-line pose initialization and real-time vehicle self-localization tasks. We denote the coordinate of the units in the sub-map M of the SGM as $m_1, ..., m_J$, and the units in the observed scan S as $s_1, ..., s_K$, where J and K are the number of units respectively. For the initialization task, the original pose must search a wide range to avoid various local minima. At the same time, the calculation can take relatively longer (several seconds is acceptable). Therefore, to keep as much map detail as possible, the SGM is in 3D formed by cubes. On the contrary, the localization task can inherit a much more accurate initial position from the previous frame while it requires strict real-time performance (typically 100ms), the SGM is in 2D formed by squares.

A. On-line Pose Initialization

In order to initialize the vehicle pose in GNSS denied areas, this paper proposes a GMM-based semantic categories to represent the pose initialization problem. We first generate a 3D SGM, and characterize each semantic category of this study as a Gaussian model. This model only focuses on the horizontal distribution as all of the three semantic categories in this study are vertically uniformly distributed. We can regard each cube of the observed scan $S_{K\times 3}$ as a mean value of the GMM, and each cube of the sub-map $M_{J\times 3}$ as the corresponding Gaussian distributed samples. The response probability of the GMM can be represented as

$$P(m_j) = \sum_{k=1}^{K} P(s_k)P(m_j|s_k) \tag{1}$$

where $P(s_k)$ is each component of the GMM of cube k. By considering the category probability of m_j and a penalty term for outliers and noise w inspired by [25], we can extend the expression as

$$\begin{aligned} P(m_j, C_{m_j}) &= w\frac{1}{J} \\ &+(1-w)\sum_{k=1}^{K} P(C_{m_j}|m_j, s_k)P(m_j|s_k)P(s_k) \end{aligned} \tag{2}$$

where C_{m_j} is the corresponding semantic category of cube m_j. We define the semantic confidence to associate the k^{th} scan cube to the j^{th} map cube as

$$P(C_{m_j}|m_j, s_k) = \begin{cases} \frac{\max(n_p,n_f,n_r)}{N} & C_{m_j} = C_{s_k} \\ 0 & C_{m_j} \neq C_{s_k} \end{cases} \tag{3}$$

where C_{s_k} is the semantic category of cube s_k of the scan, and n_p, n_f, n_r are the number of points in each semantic category of poles, facades, and road surface marks of cube m_j respectively, while N is the total point number in cube m_j. And we have

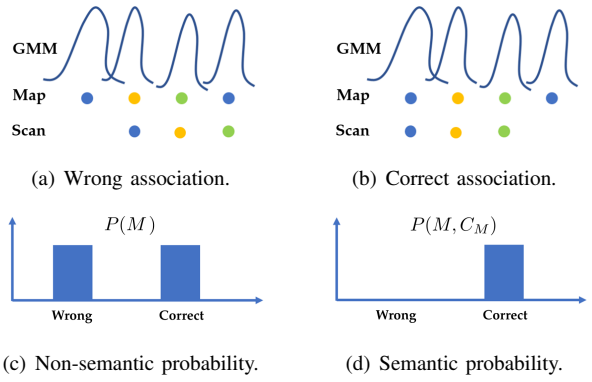
$$P(m_j|s_k) = \frac{1}{2\pi|\Sigma_k|^{\frac{1}{2}}} \exp(-\frac{1}{2}(m_j - s_k)^T\Sigma_i^{-1}(m_j - s_k)) \tag{4}$$

where Σ_k is the variance of the k^{th} component need to be solved. The pose initialization can be represented as

$$T^* = \arg\max_T P(M, C_M) = \prod_{j=1}^{J} P(m_j, C_{m_j}) \tag{5}$$

where the transformation matrix T is to decide the data association pairs of m_j and s_k in Equ. 2. T^* is to be found by maximizing the data association probability.

To solve the Σ_k and T, we use the expectation-maximization (EM) algorithm, whose solving process can be found in [25]. The role of semantic categories in this process is shown in Fig. 3. The three semantic categories are denoted as circles in blue, yellow, and green. Traditional non-semantic localization methods like CPD only consider the geometric distances between the points (or grids) between the observed scan and the map. Therefore they cannot distinguish the wrong data association (3(a)) and the correct one (3(b)).



(a) Wrong association. (b) Correct association.

(c) Non-semantic probability. (d) Semantic probability.

Fig. 3. An example of localization initialization using semantic categories in the probabilistic data association.

Considering the resolution difference between one frame LiDAR scan and the dense map, which is challenging for scan-to-map pose estimation, the semantic grid map representation can efficiently narrow such gap by down-sampling the map and enhancing the sparse scan at the same time. The GMM ensures such a strategy to reach a localization accuracy exceeding the grid resolution.

B. Real-time Localization

Probabilistic data association provides an effective framework for solving the impact of incorrect data association on the localization algorithm. As mentioned at the beginning of this section, we denote the squares in the 2D SGM as $M = \{m_j\}$, and the squares in the observed scan after gridding as $S = \{s_k\}$. The associated pairs set between M and S is denoted as $A = \{a_{j,k}\}$ where $a_{j,k} = (m_j, s_k)$. The residual error is denoted as $\varsigma = M - T \times S$ where T is the transformation matrix. The semantic category is denoted as C. The localization problem can be represented as

$$T^* = \arg\max_T P(\varsigma, C, A|M, S) \tag{6}$$

Use the Bayes Rule, this product is factored as

$$\begin{aligned} &P(\varsigma,C, A|M, S) \\ &\propto \underbrace{P(\varsigma|A, M, S)}_{error} \underbrace{P(C|A, M, S)}_{label} \underbrace{P(A|M, S)}_{geometry} \end{aligned} \tag{7}$$

The error term is defined as

$$P(\varsigma|A, M, S) = \prod \exp(\frac{-\|m_j - Ts_k\|^2}{2}) \tag{8}$$

and the label term is same to Equ. 3.

Eventually, this paper adopts the traditional geometric association as a protection term. To avoid overemphasizing the effect of Euclidean registration and thus weakening the semantic information, this paper uses the k nearest neighbors (KNN) structure under the uniform distribution to facilitate the search of the k nearest association category as

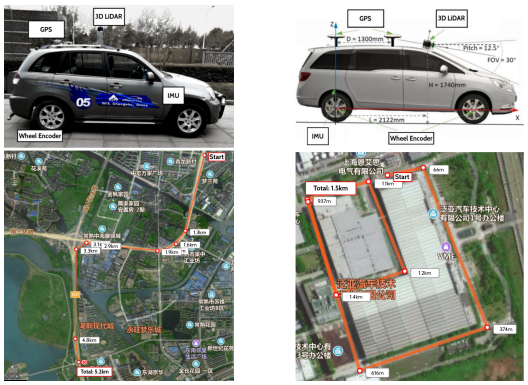
$$p(A|M, S) = \begin{cases} 1/k & knn \\ 0 & otherwise \end{cases} \tag{9}$$

According to the above association method, this paper assumes that the errors conform to the Gaussian distribution. The model needs to solve two unknown variables, one is data association probability, and the other is the pose transformation matrix T. The EM algorithm is also used to solve this problem as elucidated in reference [25].

V. Experimental Evaluations

We evaluated our method in two scenarios. The first one is an express road with heavy traffic to test the performance under strong dynamic interferences, as shown in Fig. 4(a). The second one is a factory with a generally static environment to confirm the compatibility, as shown in Fig. 4(b). The two vehicle platforms are equipped with a HESAI Pandar-40P LiDAR and a Velodyne VLP-16 LiDAR respectively. Both platforms have computing resources of the Intel i7-7567U CPU @3.5GHz with 16GB memory. The calculation times of our method in the experiments include the semantic features extraction from the observed scan, which is based on geometric rules.

In the map generation process, the GNSS positioning results are used as ground truth data. Then, the SPCM is generated from the semantically segmented point cloud consistent with GNSS [2]. Fig. 5 shows a part of the SPCM of the express road with the three semantic categories.



(a) Express road. (b) Factory.

Fig. 4. The experimental platforms and environments.

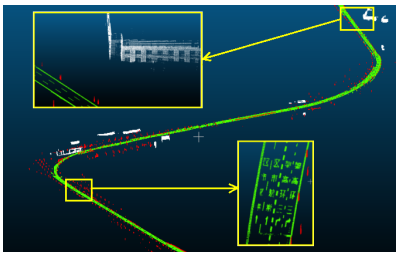


Fig. 5. An example of SPCM of the express road. The three semantic categories (poles-red, facades-white, road surface marks-green) can be easily identified.

A. On-line Pose Initialization

We randomly selected 100 different poses on the straight express road to evaluate our method. The SGM is a 3D grid map that each grid is a cube with a side length of 0.2m. As the pose initialization is sensitive to both the horizontal offset and the orientation error, in this experiment the horizontal offset is set as a uniform distribution up to 50m, and the orientation error is set to 30°, 60°, and 90°.

We compared our results with Coherent Point Drift (CPD) [25], which is a widely used method for pose initialization. The result can be found in Table I. It shows that our method has a better robustness and overall accuracy especially when the initial pose is set with the orientation error over 60°. We also compared the calculation time of orientation error at 90°which is considered as the worst case. The proposed method takes less than half time than CPD. It proves that the semantic-category-based method in this paper can speed up the iteration and reduce the time consumption.

From Table I we can see that CPD generally converge to the correct heading angle when the orientation error is set to 90°on the straight road experiment. We also demonstrate a special case of the pose initialization at a conjunction of the express road to further illustrate the effectiveness using semantic categories under the same orientation setting, as shown in Fig. 6. The semantic categories of road surface marks and poles are represented in red and green, and the observed scan is rendered in white, as shown in Fig. 6(b). Because the distance between the road surface marks of the

TABLE I

ACCURACY AND CALCULATION TIME EVALUATIONS FOR THE POSE INITIALIZATION EXPERIMENT.

	Trans.(m)	30°	60°	90°
CPD	Mean	0.17	0.19	6.42
	Max	0.18	0.50	67.3
Our method	Mean	0.08	0.18	**0.13**
	Max	0.12	0.24	**0.30**
	Yaw.(°)	30°	60°	90°
CPD	Mean	0.18	0.19	3.83
	Max	0.20	0.48	6.80
Our method	Mean	0.13	0.13	**0.11**
	Max	0.16	0.16	**0.18**
	Calculation time (s)			90°
CPD	Mean			7.25
Our method	Mean			**3.23**

TABLE II

ACCURACY EVALUATIONS FOR THE LOCALIZATION EXPERIMENT ON THE EXPRESS ROAD.

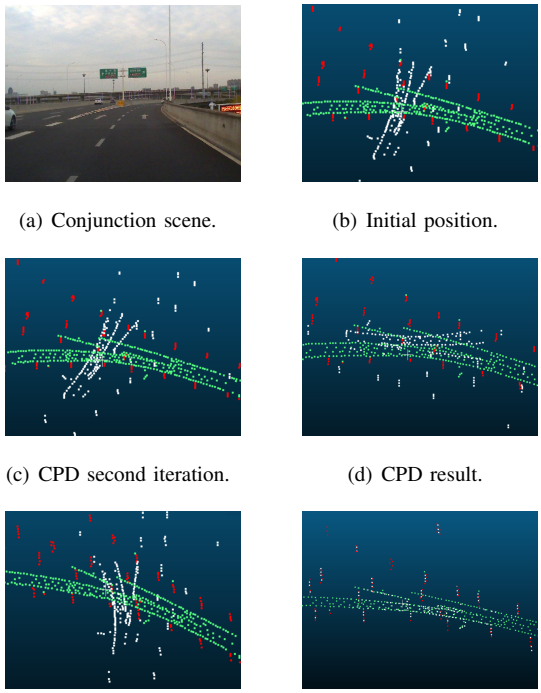
	Lat.(m)	Lon.(m)	Trans.(m)	Yaw.(°)
Semantic ICP	0.20	0.24	0.31	**0.20**
Grid Localization	0.11	≥ 2	≥ 2	≥ 2
Poles	0.37	0.33	0.55	1.86
Road marks	0.10	≥ 2	≥ 2	0.37
Facades	0.09	-	-	0.54
Our method	**0.08**	**0.12**	**0.16**	0.27

observed scan and the poles of the map are geometrically closer, CPD rotates to the wrong direction from the very beginning of the iteration, and eventually converged to the local minimum. On the contrary, the semantic category plays an important role, and leads the iteration to the correct pose.

B. Real-time Localization

In the 5.2km express road experiment, we compared our result with the Semantic ICP [23] and a traditional non-semantic approach of the occupancy grid localization using weighted point cloud[26]. We also compared our result with using each one of the three semantic categories separately to show the effectiveness using multiple semantic features.

Table II shows that our method significantly outperformed other methods and the stand-alone semantic categories in terms of transformation accuracy. For the Grid Localization method and stand-alone road surface marks, the express road is not a geometrically salient scenario in longitude (also known as corridor effect). Due to such failure, Gird Localization also failed to achieve reasonable yawing accuracy. For the same reason, the facades are parallel to the road direction; therefore, they can't provide any longitudinal pose constrain. The comparison of the calculation time proves the efficiency of this approach, as shown in Table III. Table IV compares the size of different kind of maps, from where we



(a) Conjunction scene. (b) Initial position. (c) CPD second iteration. (d) CPD result. (e) Our method second iteration. (f) Our method result.

Fig. 6. A special case at the conjunction of the express road that CPD failed to converge to the correct position.

can see proposed semantic grid map takes up the smallest storage space.

TABLE III

CALCULATION TIME EVALUATIONS FOR THE LOCALIZATION EXPERIMENT ON THE EXPRESS ROAD.

Method	Mean operation time(ms)
Semantic ICP	150.40
Grid Localization	44.03
Poles	15.14
Road marks	16.34
Facades	14.08
Our method	23.41

In the factory experiment, our method can also achieve a comparable accuracy to the express road condition, where the comparison can be found in Table V. The main reason is that in factory experiment, there are more structural features which are beneficial for SGM-based localization.

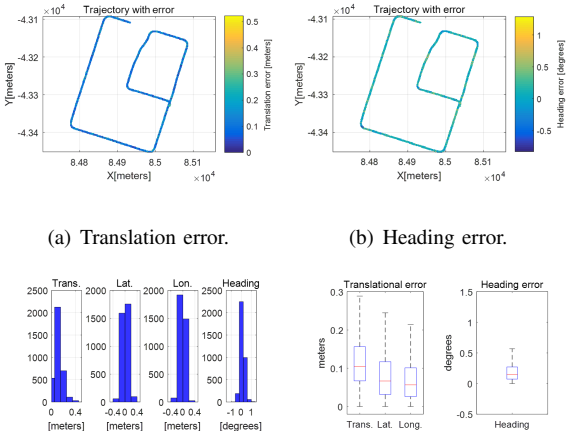
VI. CONCLUSION

In this paper, we proposed a localization method based on the semantic grid map (SGM) with poles, facades, and road surface marks. Such map is small in size and rich in information. By introducing the Gaussian mixture model (GMM) to the semantic features, the corresponding pose initialization method improved the robustness and accuracy while reduced the calculation time by half comparing to the traditional non-semantic baseline. In the real-time localization process, this

TABLE IV

MAP STORAGE SIZE COMPARISON.

Map structure	Size(MB/km)
Point cloud map	$\geq$ 1000
Semantic point cloud map	34
Grid map	5.3
Semantic grid map (Ours)	**1.1**



(a) Translation error. (b) Heading error.

(c) Error histogram. (d) Error boxplot.

Fig. 7. Localization result in the factory environment.

paper introduced grid probability to implement a new data association strategy with semantic information. Experimental results show that our proposed method is robust and accurate in not only dynamic scenarios, but also static environments which guaranteed the adaptability.

TABLE V

ACCURACY COMPARISON BETWEEN THE DYNAMIC SCENES AND STATIC SCENES FOR THE LOCALIZATION EXPERIMENT.

	Trans.(m)	Lat.(m)	Lon.(m)	Yaw.(°)
Express road	0.14	0.06	0.11	0.21
Factory	0.12	0.07	0.08	0.19

REFERENCES

[1] F. Yu, J. Xiao, and T. Funkhouser, "Semantic alignment of lidar data at city scale," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, United states: IEEE, 2015, pp. 1722–1731.

[2] Y. Chen, M. Yang, C. Wang, and B. Wang, "3d semantic modelling with label correction for extensive outdoor scene," in *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2019, pp. 1262–1267.

[3] P. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[4] M. Magnusson, "The three-dimensional normal-distributions transform: an efficient representation for registration, surface analysis, and loop detection," Ph.D. dissertation, Örebro universitet, Stockholm, 2009.

[5] J.-E. Deschaud, "Imls-slam: Scan-to-model matching based on 3d data," in *2018 IEEE International Conference on Robotics and Automation(ICRA)*. Brisbane, QLD, Australia: IEEE, 05 2018, pp. 2480–2485.

[6] J. Levinson and S. Thrun, "Robust vehicle localization in urban environments using probabilistic maps," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, 2010, pp. 4372–4378.

[7] G. Wan, X. Yang, R. Cai, H. Li, Y. Zhou, H. Wang, and S. Song, "Robust and precise vehicle localization based on multi-sensor fusion in diverse city scenes," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. Brisbane, QLD, Australia: IEEE, 2018, pp. 4670–4677.

[8] J. Yang, H. Li, and Y. Jia, "Go-icp: Solving 3d registration efficiently and globally optimally," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1457–1464.

[9] F. Pomerleau, F. Colas, R. Siegwart *et al.*, "A review of point cloud registration algorithms for mobile robotics," *Foundations and Trends® in Robotics*, vol. 4, no. 1, pp. 1–104, 2015.

[10] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp." in *Robotics: science and systems*, vol. 2, no. 4. Seattle, WA, 2009, p. 435.

[11] J. Serafin and G. Grisetti, "Nicp: Dense normal based point cloud registration," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015, pp. 742–749.

[12] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Autonomous Robots*, vol. 41, no. 2, pp. 401–416, 2017.

[13] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Madrid, Spain: IEEE, Oct 2018, pp. 4758–4765.

[14] R. Dubé, D. Dugas, E. Stumm, J. Nieto, R. Siegwart, and C. Cadena, "Segmatch: Segment based place recognition in 3d point clouds," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. Singapore, Singapore: IEEE, 2017, pp. 5266–5272.

[15] W. Lu, Y. Zhou, G. Wan, S. Hou, and S. Song, "L3-net: Towards learning based lidar localization for autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Long Beach, CA, USA: IEEE, 2019, pp. 6389–6398.

[16] D. P. Paudel, A. Habed, and L. Van Gool, "Optimal transformation estimation with semantic cues," in *2017 IEEE International Conference on Computer Vision (ICCV)*. IEEE, 2017, pp. 4668–4677.

[17] R. Spangenberg, D. Goehring, and R. Rojas, "Pole-based localization for autonomous vehicles in urban scenarios," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Daejeon, Korea: IEEE, 2016, pp. 2161–2166.

[18] M. Sefati, M. Daum, B. Sondermann, K. D. Kreisköther, and A. Kampker, "Improving vehicle localization using semantic and pole-like landmarks," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. Redondo Beach, CA, United states: IEEE, 2017, pp. 13–19.

[19] A. Kampker, J. Hatzenbuehler, L. Klein, M. Sefati, K. D. Kreiskoether, and D. Gert, "Concept study for vehicle self-localization using neural networks for detection of pole-like landmarks," in *International Conference on Intelligent Autonomous Systems*. Baden-Baden, Germany: Springer, 2018, pp. 689–705.

[20] F. Poggenhans, N. Salscheider, and C. Stiller, "Precise localization in high-definition road maps for urban regions," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Madrid, Spain: IEEE, 10 2018, pp. 2167–2174.

[21] A. Hata and D. Wolf, "Road marking detection using lidar reflective intensity data and its application to vehicle localization," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. Qingdao, China: IEEE, 2014, pp. 584–589.

[22] J. K. Suhr, J. Jang, D. Min, and H. G. Jung, "Sensor fusion-based low-cost vehicle localization system for complex urban environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1078–1086, 2016.

[23] S. A. Parkison, L. Gan, M. G. Jadidi, and R. M. Eustice, "Semantic iterative closest point through expectation-maximization." in *BMVC*, 2018, p. 280.

[24] X. Chen, A. Milioto, E. Palazzolo, P. Giguère, J. Behley, and C. Stachniss, "Suma++: Efficient lidar-based semantic slam," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 4530–4537.

[25] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[26] L. Guo, M. Yang, B. Wang, and C. Wang, "Occupancy grid based urban localization using weighted point cloud," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. Rio de Janeiro, Brazil: IEEE, 2016, pp. 60–65.