

Detection and Classification of Pole-like Landmarks for Domain-invariant 3D Point Cloud Map Matching

Sun Yifei
University of Fukui
Fukui, Japan

mf210202@g.u-fukui.ac.jp

Li Dingrui
University of Fukui
Fukui, Japan
mf230323@g.u-fukui.ac.jp

Ye Mingyong
University of Fukui
Fukui, Japan
yymm22808@gmail.com

Tanaka kanji
University of Fukui
Fukui, Japan
tnkknj@u-fukui.ac.jp

Abstract— *In 3D point cloud-based visual self-localization, pole landmarks have a great potential as landmarks for accurate and reliable localization due to their long-term stability under seasonal and weather changes. In this study, we aim to explore the use of recently developed deep learning models for pole classification in the context of pole landmark-based self-localization. Specifically, the proposed scheme consists of two main modules: pole map matching and pole class matching. In the former module, local pole map is constructed and its configuration is compared against a precomputed global pole map. An efficient RANSAC map matching is employed to achieve a good tradeoff between computational efficiency and accuracy. In the latter pole class matching module, the local and global poles paired by the RANSAC map-matching are further compared by means of pole attribute class. To this end, a predefined set of pseudo pole classes is learned via k-means clustering in a self-supervised manner. Experiments using publicly available NCLT dataset showed that the pole-like landmark classification method has an improved effect on the visual self-localization system compared with the baseline method.*

Keywords— *RANSAC Map Matching, Point cloud, Pole detection, Pole classification, K-means clustering*

I. INTRODUCTION

Accurate and reliable localization and map building in the face of various outdoor environments is a fundamental problem for mobile robots and autonomous vehicles. At present, the positioning system based on global satellite navigation system is easily affected by objective conditions such as terrain and weather because the positioning accuracy of the satellite system are often unreliable due to confusing signals and satellite occlusions. Positioning based on dense maps such as grid maps, point clouds or polygon grids is more demanding on memory for negative, dense maps and cannot

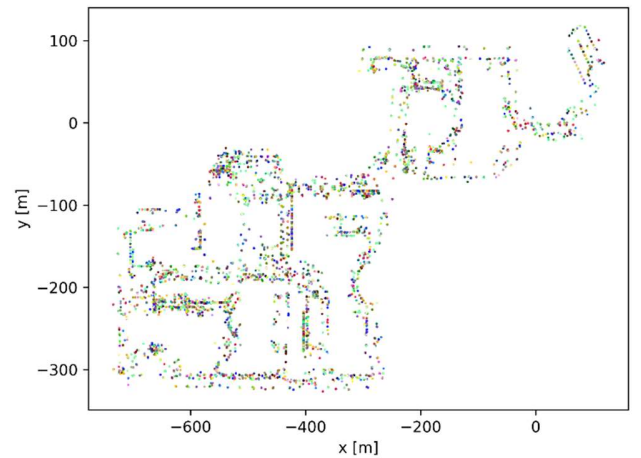


Figure 1: Pole landmark map generated from NCLT dataset. Different color corresponds to different pole class.

cope with larger scale map-based self-localization.

Furthermore, landmark maps can reduce memory occupation by several orders of magnitude while ensuring positioning accuracy.

Vision is the primary sensor for many localization and place recognition algorithms. We chose a method based on visual self-localization of polar landmarks in a 3D point cloud that relies on polar landmarks extracted from mobile LiDAR data. Pole landmarks appear as part of street lights, traffic signs and tree trunks. In addition, pole landmarks are widespread and abundantly distributed in urban areas. It is stable, stable over time under seasonal and weather changes, and its geometry is well defined. Therefore, if a static object can be found as a positioning reference in a dynamic environment, the positioning accuracy will be greatly

improved.

3D LiDAR data is superior to 2D image data in dealing with problems such as illumination and background confusion, and the application of 3D LiDAR data is starting to gain popularity. The localizer proposed by Kim et al. [1] learns point cloud descriptors called Scan Context Images (SCI) and performs robot localization on a grid map by formulating the place recognition problem as place classification using convolutional neural networks. In Cabo et al. [2], Cuenca et al. [3], Yu et al. [4], Zheng et al. [5] and Wu et al. [6], excellent research results were achieved using high-density and high-precision LiDAR data acquired by MLS (Mobile Laser Scanner) to extract urban street facilities. In Schaefer et al. [7] the authors used a visual navigation system with LiDAR to generate environment maps and then extracted pole-like landmarks from the maps. Finally, the effectiveness of global localization is demonstrated by matching the current scan with the extracted landmarks.

Our observations are that pole landmarks trade discriminative power for detection robustness. A pole is represented as a 2D point on the map. Such low-dimensional representations often make it difficult to distinguish similarly positioned poles. We attach the minimum necessary attribute information to the pole. By doing so, we aim to dramatically improve the identifiability while maintaining the robustness and lightness of the pole landmark. Specifically, this attribution task is formulated as the problem of classifying a pole's attributes into one of a predefined set of attribute candidates (Figure 1).

The proposed scheme consists of two main modules: pole map matching and pole class matching. In the former module, local pole map is constructed and its configuration is compared against a precomputed global pole map. An efficient RANSAC map matching is employed to achieve a good tradeoff between computational efficiency and accuracy. In the latter pole class matching module, the local and global poles paired by the RANSAC map-matching are further compared by means of pole attribute class. To this end, a predefined set of pseudo pole classes is learned via k-means clustering in a self-supervised manner. In addition, we establish a baseline method for generating map hypotheses through pole-based self-localization only, and match and score the two sets of hypotheses generated to obtain matching accuracy. Experiments using publicly available NCLT dataset showed that the pole-like landmark classification method has

an improved effect on the visual self-localization system compared with the baseline method.

II. POLE CLASS MATCHING

A. Pole Landmark Identification

The stability of pole-like landmarks under different seasonal and weather conditions has become a quality landmark for self-localization series studies. In recent years, many authors have addressed specific problems of vehicle localization using pole landmarks extracted from 3D LiDAR scans. Pole extraction is an important component of localization systems, where researchers are interested not only in fitting geometric primitives to the data and determining pole coordinates, but also in accurate segmentation of pole-like landmarks. All methods provide pole extraction systems, and some researchers have proposed composite pole localization systems. Most of the solutions offered in the study include at least the following two components: a pole detector and a landmark-based pose estimator. A complete localization system is formed by the pole extraction module and the landmark localization module. The detector developed by Li [8] et al. was used to segment the space around the LIDAR sensor and calculate the laser reflection count for each voxel. The poles are then assumed to be located at points that exceed the reflection count threshold.

B. Pole Landmark Extraction

The principle of laser distance measurement is based on focused laser imaging and measuring the return time of the laser signal to calculate distances using sensors and digital acquisition devices. When dealing with distant objects, the laser can be adapted to provide imaging and optical signal information by providing a specific wavelength. As the technology continued to advance, it eventually led to the creation of LIDAR (Light Detection and Ranging). As the technology continues to evolve, LIDAR is now quite mature, and numerous LIDAR companies have tested and commercialized it in self-driving cars, advanced driver assistance systems, mapping, robotics, infrastructure and smart city applications.

Schaeffer et al. proposed a method to subdivide space and extract poles based on the number of laser reflections in each voxel. The occupied and available space is explicitly modeled by the start and end points of the scan. The pole extraction module proposed by Schaefer takes as input a set of

registered 3D LiDAR scans, allows estimating the width properties of the poles, and outputs the 2D coordinates of the detected pole centers with respect to the ground plane. The pole extraction module requires the construction of a 3D occupancy map of the scan space, and the pole feature detector can regress the generated pole map into a set of detected pole position and width estimates for each voxel.

C. Pole Landmark Clustering

The standard k-means clustering algorithm is described as follows. The computation begins with randomly selecting $k=200$ observations from the data set and used as initial means. Scattering the initial means and setting up random partitions to place all initial means close to the center of the dataset, the algorithm is performed by alternating between two steps [9].

Assignment step: assign each observation to the cluster with the closest mean: the cluster with the least square Euclidean distance.

$$S_i^{(t)} = \{x_p: ||x_p - m_i^{(t)}||^2 \leq ||x_p - m_j^{(t)}||^2 \forall j, 1 \leq j \leq k\}$$

where each is assigned to exactly one, even though it can be assigned to two or more of them.

Update step: recalculate the mean (center of mass) of the observations assigned to each cluster.

$$m_i^{(t+1)} = \frac{1}{S_i^{(t)} \vee \sum_{x_j \in S_i^{(t)}} x_j}$$

The computation is considered converged and terminated when the allocation no longer changes [10].

III. POLE MAP MATCHING

Interest in the use of mobile geographic information systems (GIS) in a variety of automotive applications has increased dramatically over the past few years. Car navigation systems use digital maps to guide drivers to their desired destinations. The next generation of driver assistance systems will use enhanced maps to provide precise navigation cues (including locations of speed limits, gas stations, restaurants, etc.) and to assist with vehicle control. In addition, methods to autonomously augment existing maps using video and LiDAR sensors are being developed. The measured vehicle locations are robustly assigned to road segments in the digital map. This process is called map matching. Motion structure has long been a very active area of research in computer vision. The

estimation of known structures and motions is highly sensitive to noise and outliers. Therefore, during map matching, it is important to minimize the cost function as close as possible to the true noise and outlier conditions. The random sample consensus (RANSAC) algorithm is a widely used robust estimator that has become the dominant stochastic model validation strategy in computer vision. RANSAC and related hypothesis and validation methods have been applied to map matching, visual place recognition, and many other related studies.

A. RANSAC Algorithm

The main task of RANSAC is to generate hypotheses and score them. Among them, for the running process of RANSAC in the baseline self-localization system, n poles are selected as input, the distance between any two input poles is calculated to bring into a pre-generated distance lookup table to find a hypothesis that meets the conditions, and after rotating the hypothesis by translational transformation, if any point in the hypothesis matches the position of a point in the global map, the score is increased by 1 point, the cycle iterates through all hypotheses, and the score in the hypothesis with the highest rotation-translation matrix will be selected as the final hypothesis. For the pole-based self-localization system, the location information of the points in the hypothesis in the scoring benchmark is combined with the classification information of the pole-like landmarks. If any point in the hypothesis matches a point position in the global map, or any point with the same pole classification category c as a point classification category c in the global map, then the score is increased by 1, iterating through all hypothesis cycles, the rotation-translation matrix with the highest score in the hypothesis is selected as the final hypothesis for the pole class group. The specific flow chart is shown in Figure 2.

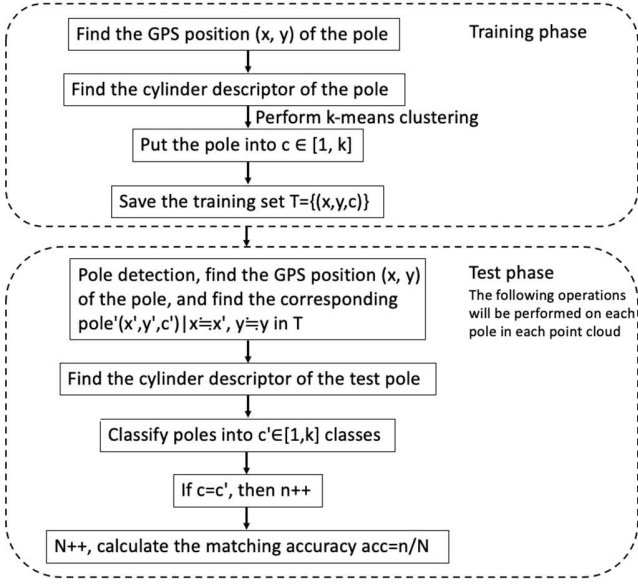


Figure 2: RANSAC Generating Hypothesis and Scoring Flowchart

IV. EXPERIMENTS

A. Dataset

We use the NCLT dataset to evaluate the proposed approach: the accuracy of the self-localization system is improved by the joint combination of pole classification and self-localization. The NCLT (North Campus Long Term) [11] dataset was acquired from two rounds of Segway robots provided by the University of Michigan campus (Figure 2). The NCLT dataset can be used for tasks such as navigation and mapping in changing environments using vision or LiDAR, and is also well suited for robotics research testing large-scale, long-term autonomous problems such as long-term localization in urban environments.

These datasets explore the campus through iterations that include seasonal changes, different weather conditions (e.g., falling leaves and rain and snow), different times of day, indoor and outdoor environments, a large number of static objects (trees, street signs, light poles, etc.), and moving things (e.g., pedestrians, bicycles, and cars), as well as long-term structural changes resulting from two large construction projects that are under constant construction. The pole landmarks in the NCLT dataset can serve as a good reference for the visual localization of poles in this study. Although the trajectories vary considerably between sessions, there is significant overlap in the trajectories. For more detailed information on the NCLT dataset, see [11].

B. Results

We record the results of different models on different test datasets and mathematically analyze all the results. For pole classification, the setting of variable k value in K-means algorithm, we calculated the optimal k value for the correct rate of visual position recognition system by repeated experiments.

We use two different error thresholds: 5 meters and 1 meter. When the error between the predicted positioning (x, y) and the real positioning (x, y) is smaller than the threshold, the self-localization is regarded as successful. We computed the percentage of accuracy, which is the number of successful test images normalized by the total number of test images.

After that, we tested other datasets and obtained the pole classification accuracy of the proposed and baseline self-localization frameworks as the test results. The results of the pole classification accuracy with the proposed framework and the accuracy of the baseline framework are shown in Table 1.

Test dataset	Accuracy			
	Threshold: 5 m		Threshold: 1 m	
	Ours	Baseline	Ours	Baseline
2012-03-17	97.66%	92.19%	55.69%	47.51%
2012-05-11	97.93%	91.37%	48.65%	37.16%
2012-11-04	96.45%	88.39%	51.08%	43.64%

Table 1: Performance results.

A visual aerial view of the matching of the pole classification with the proposed and baseline self-localization frameworks are shown in Figure 2. In Figure 2, the red line is the original trajectory localization point and the blue line is the predicted localization point. If the localization is successful, the points of the two colors are close to each other. The higher the localization accuracy, the higher the overlap between the red and blue lines. By presenting such a visualization, readers can visually judge the success rate of self-localization.

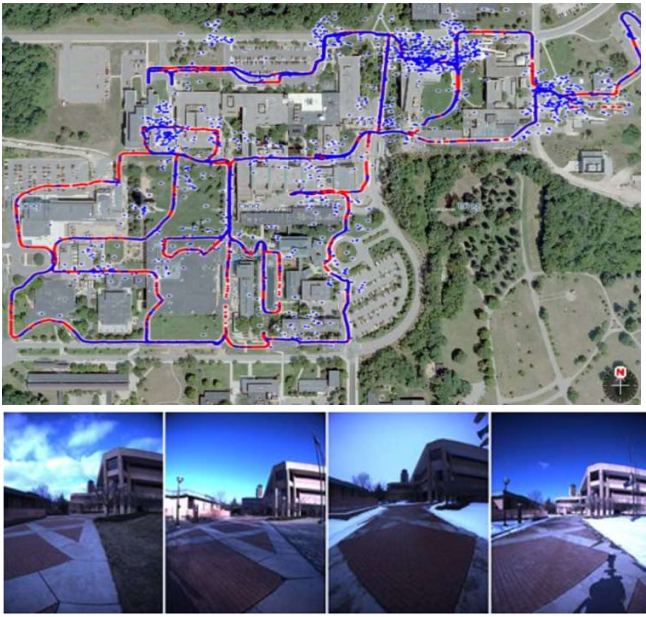


Figure 2: Experimental environment with example results of pole classification (dataset: 2012-03-17).

It is clear from the results that our proposed VPR model-based pole classification outperforms the baseline model in terms of position recognition accuracy and overall model robustness, which validates the gain of the pole-like classification-based position recognition system.

V. CONCLUSION

In this paper, we propose a new method to obtain point cloud information using a pole extractor, perform a columnar representation of the point cloud information features and classify the pole-like landmarks using a K-means clustering algorithm, and then use the RANSAC model as a decision module to generate matching hypotheses and set a baseline method as a comparison to achieve accurate and effective visual self-localization.

We use the method of dividing the LiDAR point cloud into cylindrical partitions to generate more balanced points for the cylindrical partitions, and then obtain pole-like landmark features based on the partitions to reduce the dimension of the input data, and the subsequent classification difficulty. We use K-means clustering algorithm to classify the obtained pole landmark features. The algorithm principle and process of RANSAC is described in detail using the robust RANSAC parameter estimation algorithm, and the algorithm is applied to the RANSAC generation assumptions and evaluation in the process of visual self-localization research.

We select the NCLT dataset to verify the feasibility and effectiveness of pole-like landmark-based classification in visual place recognition. The stability of the results from multiple test sets is experimentally analyzed by comparing with the baseline method and tuning the parameters with accuracy as the performance metric. The experimental data show that our proposed model achieves excellent performance in terms of self-localization accuracy and overall model robustness compared to the baseline method.

References

- [1] Kim G, Park B, Kim A. 1-day learning, 1-year localization: Long-term lidar localization using scan context image[J]. *IEEE Robotics and Automation Letters*, 2019, 4(2): 1948-1955.
- [2] Cabo C, Ordoñez C, García-Cortés S, et al. An algorithm for automatic detection of pole-like street furniture objects from Mobile Laser Scanner point clouds[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 87: 47-56.
- [3] Rodríguez-Cuenca B, García-Cortés S, Ordóñez C, et al. Automatic detection and classification of pole-like objects in urban point cloud data using an anomaly detection algorithm[J]. *Remote Sensing*, 2015, 7(10): 12680-12703.
- [4] Yu Y, Li J, Guan H, et al. Semiautomated extraction of street light poles from mobile LiDAR point-clouds[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 53(3): 1374-1386.
- [5] Zheng H, Wang R, Xu S. Recognizing street lighting poles from mobile LiDAR data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 55(1): 407-420.
- [6] Wu F, Wen C, Guo Y, et al. Rapid localization and extraction of street light poles in mobile LiDAR point clouds: A supervoxel-based approach[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 18(2): 292-305.
- [7] Schaefer A, Büscher D, Vertens J, et al. Long-term urban vehicle localization using pole landmarks extracted from 3-D lidar scans[C]//2019 European Conference on Mobile Robots (ECMR). *IEEE*, 2019: 1-7.
- [8] Li F, Oude Elberink S, Vosselman G. Pole-like road furniture detection and decomposition in mobile laser scanning data based on spatial relations[J]. *Remote sensing*, 2018, 10(4): 531.
- [9] MacKay D J C. Information theory, inference and learning algorithms[M]. Cambridge university press, 2003.

[10] Hartigan J A, Wong M A. Algorithm AS 136: A k-means clustering algorithm[J]. Journal of the royal statistical society. series c (applied statistics), 1979, 28(1): 100-108.

[11] Carlevaris-Bianco N, Ushani A K, Eustice R M. University of Michigan North Campus long-term vision and lidar dataset[J]. The International Journal of Robotics Research, 2016, 35(9): 1023-1035.