Avoiding to face the challenges of visual place recognition

Mihankhah, Ehsan; Wang, Danwei

2019

Mihankhah, E., & Wang, D. (2018). Avoiding to face the challenges of visual place recognition. Intelligent Systems Conference 2018 (IntelliSys 2018).
doi:10.1007/978-3-030-01054-6_52

<https://hdl.handle.net/10356/88443>

https://doi.org/10.1007/978-3-030-01054-6_52

© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at:
https://doi.org/10.1007/978-3-030-01054-6_52

Downloaded on 26 Apr 2023 18:20:37 SGT

Avoiding to Face The Challenges of Visual Place Recognition

Ehsan Mihnkhah and Danwei Wang
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore

Abstract—through this paper, bottlenecks of conventional place recognition techniques are studied, and a replacement strategy is proposed for each item. Conventional place recognition algorithms are extensions of object recognition techniques applied to larger scale targets known as the place landmarks. The discussion presented in this paper aims to address the challenges of detection and recognition of the places, which make this topic distinctive from detection and recognition of the objects and landmarks. The challenges are listed under related categories. The table of challenges, reasons, and the recommendations to avoid these situations is presented as the guideline for selection of proper tools for place recognition purpose.

Keywords—place detection; place recognition; non-feature-based place recognition; loop closure detection.

I. INTRODUCTION

Location awareness is a critical requirement in autonomous mobile robotics. Path planning, navigation, mission planning, and many other functionalities are highly dependent on location awareness. To fulfill different objectives, location of the robot should be known with multiple levels of precision. For example, metrically precise coordinate position estimation is required for local path planning and navigation, which is mostly addressed through SLAM (simultaneous localization and mapping) techniques [1], locally and for a limited duration. However, mission planning requires general estimation of robot location in the studied environment, which is mostly addressed through place recognition techniques [2]. Although the focus of this paper is on the latter case, recognition of previously visited places facilitates loop-closure and map-merging phase of the former group of localization techniques, to help them maintain the acceptable precision and consistency in a longer run.

Localization of mobile robots can be achieved either by external referencing [3], or through processing the on-board sensor data [4]. In the former approach, beacons and cameras can be used for precise indoor applications [5]. Moreover, sound signatures, signal strength and availability of WiFi access points, and other place-specific factors can be used for fingerprint-based location estimation in indoor applications [6]. For outdoor, GNSS readings (which is perfect under ideal satellite coverage) [7], and position calculation through distance estimation from nearby mobile phone towers (which results to a large location uncertainty radius) [8], are possible techniques in accordance with the former approach. On the other hand, the latter approach has been addressed by researchers who extended

the scope of metrically accurate localization through incremental map generation in outdoor environment [9]. Other than that, the image retrieval techniques that correspond an observation to one place out of a group of expected places, form another major direction to address place recognition problem [10]. These image retrieval techniques for place recognition are mainly the extensions of object recognition techniques. While localization through external referencing requires minimal processing, indoor solutions cannot be extended to outdoor, and outdoor solutions lose coverage in indoor environment [11]. On the other hand, place recognition using onboard equipment confront limitations such as drift and sensitivity to environmental conditions [2]. This paper addresses the limitations of place recognition through processing the onboard equipment, and proposes alternative solutions to avoid facing such limitations. A place recognition procedure is consisted of several steps. Each step is a major component of the methodology that substantially affects the recognition results. This paper presents a guideline for selection of proper components for place recognition solutions. The focus of the paper is on the philosophy, including definition, sensing, and procedure steps, rather than the analysis of the results of numeric case-studies.

Section II of this paper studies the existing approaches to place detection. Section III explains the shortcomings of the common approaches by studying the definition, sensing strategies, and the procedure steps (place recognition components). Proposed amendments are suggested in section III as well. In section IV the paper is summarizes through presentation of the guideline table, and the conclusion is made successively.

II. EXISTING PLACE RECOGNITION TECHNIQUES

In this section, two major conventional approaches for place recognition, which only use the onboard equipment, are studied.

A. Techniques inspired by SLAM

SLAM has been an ongoing research for the past few decades [12-14]. Perfect SLAM algorithms were proposed in the literature and mature implementations are available as commercial products [15], and as open-source solutions [16]. Some SLAM algorithms compare the consequent observations point-by-point [17], and the others only process the feature-points [18]. While pointwise comparison is less sensitive to the shape of the objects that the observation point is surrounded by,

feature-wise comparison is less demanding in terms of computation and storage[19], which are the two factors that form valid concerns in large-scale missions and in the long run.

According to the map, when metrically accurate current coordinate location of the mobile robot is close to a previously explored point, the explored area around the current position is recognized as a revisited place. However, there is an estimation error in each incremental step of map generation that propagates through time, and results to a drifted map, which ultimately, results to inaccurate position estimation. SLAM has negligible limitations for 2D indoor applications. Recent LiDAR-based and camera –based indoor 3D SLAM techniques have also proven acceptable consistency for indoor environment [20]. However, maintaining a consistent 2D or 3D SLAM, as a general solution for large-scale outdoor environment for long-term autonomy, is still an open challenge.

One suggested solution to overcome this limitation is using the offline generated map of the environment. One such implementation is NDT¹ matching [21], which is also used to localize a driverless car on the road [22]. In this implementation, the 3D point cloud of the studied place is stored as the known map of the environment. LiDAR inputs are used to match the current observation with the offline map, and initialization is handled through GNSS.

Another general treatment is to reset the online map generation drifts through GNSS readings (under ideal satellite coverage) or environment signal-finger-printing (in presence of the necessary infrastructure) [23]. However, this approach is very much dependent on external referencing, which is beyond the scope of this paper,

In SLAM-based localization, if a revisited place is observed frequently, to the extent that the drift falls within the scan matching tolerance, small drifts can be compensated, and the map accuracy can be constantly recovered, and so the location estimation precision can be maintained. However, if the revisited place is not observed for a long time, to the extent that the accumulated drift disallows matching the new observation with the correct piece of the map, it becomes very challenging to recognize a revisited place. To recover from this situation, recognition of revisited places should be handled through a technique that is independent from SLAM.

B. Techniques inspired by image retrieval

The most commonly applied technique for place recognition, which is generally independent from SLAM, is image retrieval. In this technique, current observation is compared against a database of stored images to recognize the best match. Each image in the database is associated to a place. Therefore, matching the current observation with a previously stored frame conveys the similarity to the place with which it is associated with [2].

Observations can be 2D images (colour intensity information) [24], gray 3D point clouds (depth information) [25-

27], or RGBD images (both colour intensity and depth information) [28, 29].

Image retrieval techniques used for place recognition are extended versions of object recognition algorithms. Similar to object recognition methodologies, which identify each object through a set of specific features, in place recognition methodologies based on image retrieval, places are either recognized as a result of recognition of an outstanding landmark or through recognition of a set of specific type of features [10, 30]. The procedure involves the phase of breaking the images into smaller patches, known as segmentation step. Followed by that, features are extracted from the patches. Then features are converted to numeric structures, known as feature descriptor vectors. Each image is described through a set of feature-descriptors. The set of feature descriptors is called the image descriptor. Feature descriptors are stored in a database. For each new observation, feature descriptors are computed and compared against the stored feature-descriptors in the database. The most similar stored image to the current observation is the one that embodies the highest number of similar feature-descriptors.

Feature selection is a critical step in image retrieval. Invariance of features is ideally expected against scaling, rotation, shifting, and deformations [31]. Enough number of features should be consistently identifiable to compute a representative image descriptor. Several techniques are available for feature detection. Working on camera frames, simplest visual features are edges [32] and corners [33]. For example, the Hessian feature detector computes the Hessian for points that exhibit high derivative values in two orthogonal directions [34], a point is selected as a feature when determinant of the Hessian becomes maximum at that point. The Harris detector [33] is a geometric feature detector that finds corner-like structures from image regions where the second moment matrix has large eigenvalues. LoG² [35], and a close approximation of it, DoG³ [36], are also well-known feature detectors. A more resilient detector against rotation, scaling, illumination changes, and input noise [37] is Harris-Laplacian detector [38] that combines the Harris corner detection with LoG. If the Harris-Laplace is applied to Hessian, the result is called Hessian-Laplace detector [39].

When input is in form of binary point cloud data, feature extraction in 3D case can be achieved through processing the meshes [40]. For example, in [40], features points are extracted when second difference operator is applied to mesh edges. Likewise, the work [41] extracts feature lines from meshes by applying third order derivatives. Work [42] extracts curvature extremums as feature points. The work [43] uses region growing method, which is a meshless feature detection methodology that identifies regions with sharp features from point cloud segments. Feature lines are found subsequently. Through a statistical approach, PCA⁴ extracts salient features from point clouds. Extracted features are the direction of principle axes in the space along which the studied data points have the highest variance,

1 - Normal Distributions Transform

2 - Laplacian-of-Gaussian

3 - difference-of-Gaussian

4 - Principle Component Analysis

and the lowest variance [44]. This method can be applied to multi-dimensional data for place recognition[45]. In a recent work [46], 3D point clouds are projected to different planes in the space, and similar feature detector is used to find features holding statistic information about the studied point cloud.

Majority of place detection techniques use the SIFT⁵ descriptor [4]. SIFT algorithm calculates the local gradient magnitudes and their orientations for each key-point. Subsequently, it calculates the orientation histograms for sub-regions. The SIFT descriptor is a vector that holds values of all orientation histograms for each key-point. An extension of SIFT is called GLOH⁶ descriptor that is designed to enhance robustness and distinctiveness of description [47]. SURF⁷ descriptor [48] is a well-known computationally efficient replacement to SIFT. [37]. “Haar wavelets” [49] are used in SURF. The SURF algorithm generates a structure that embeds intensity statistics along 2D coordinate directions. There are other types of descriptors that work on content distribution. For example, geometric histograms [50] or shape context [51] are used to form image descriptors. Work [52] uses 3D sparse visual features, which are mainly the corners, introduced in [53]. In [52], description procedure is initiated by detection of landmarks followed by computation of 3D Gestalt feature descriptor [54-56]. Work [57], introduces surface entropy features (SURE) [58] for place recognition. SURE features embed the statistics of the distribution of local surface normal vectors. Invariant feature selection is an ongoing challenge in place recognition.

Using the bag-of-words model [59] is the most referenced methodology for storage, retrieval and comparison of feature descriptors [60]. Bag-of-words model associates feature descriptor to words. Subsequently, word processing techniques are used to compare descriptor structures. Bag-of-words model discards the geometric and spatial relevance among image features. Since bag-of-words dictionary can get large in the long run, optimized data access techniques were suggested in the literature to handle this problem [61].

As explained earlier, snapshot pairing through viewpoint tolerant feature matching is the common approach in place recognition based on image retrieval [62]. Combination of features and feature descriptors have been used in place recognition research. These descriptors were stored, retrieved, and compared through modified versions of the bag-of-words model.

III. CHALLENGES OF EXISTING METHODOLOGIES AND THE PROPOSED REPLACEMENTS

A. SLAM specific challenges

As explained in section II.A, SLAM solves place recognition problem for short-term autonomy, in small-scale indoor environments. However, for long-term autonomy, and large-scale outdoor environment, SLAM requires an independent mechanism to handle recognition of revisited places to cancel the drifts (which is known as the problem of “closing the loop” in SLAM community). Therefore, recognition of places through SLAM, for large-scale outdoor environment in long-term

autonomy, is a mutually dependent problem. Consequently, for long-term autonomy in large-scale outdoor environment, it is necessary to solve SLAM and place recognition problems individually and independently.

Source of inconsistency in SLAM algorithms is the accumulation of error, caused by incremental map generation and location estimation through the comparison of consecutive observations. Therefore, the source of problem is the incremental computation approach, which embeds an intrinsic integrator.

SLAM is prone to facing storage and computational limitations in the long run, when pointwise scan matching is involved. Therefore, for large-scale long-term applications, pointwise methods should be avoided. On the other hand, feature-based methods are successful whenever enough number of expected type of features are guaranteed to exist consistently in all observations. Invariant feature selection and feature association are major challenges in feature-based SLAM.

Therefore, for large-scale long-term autonomy in outdoor missions, or even for combined indoor and outdoor missions, SLAM-based place recognition should be complemented with external referencing. However, if the presence of external referencing is assumed, place recognition can be handled solely through the external referencing equipment, and without any reliance on SLAM.

B. Challenges of place recognition through common image retrieval techniques

As explained in section II.B, place recognition based on image retrieval is commonly handled through landmark recognition and feature matching techniques.

Although recognition of outstanding landmarks is a concrete proof for recognition of a place, not all places contain an outstanding landmark. Therefore, this approach is limited only to recognition of places that include outstanding landmarks. In another words, mapping a place to a landmark is not generally a valid assumption for recognition of all places.

Even if features are directly extracted from observations without prior landmark recognition, existence of enough number of specific type of features in every observation is a limiting assumption. This problem enroots in invariant and consistent features selection challenge. This limitation advocates incorporation of non-feature-based description methods such as [63, 64].

Under the assumption of existence of enough number of feature points, the philosophy of a place being represented only through limited number of feature-points is a matter of dispute. The assumption of a place being represented through limited key-points is resulted by extending object recognition philosophy to place recognition. The assumption of an object being fully represented through limited features is logical. However, every single piece of the surrounding of a place participate in characterization of that place. This necessitates

5 - Scale Invariant Feature Transform

6 - Gradient location-orientation histogram

7 - Speeded-Up Robust Features Descriptor

incorporation of holistic description techniques such as [65] rather than local ones that operate only on few image patches.

Bag-of-words database is reported to have included over 100 million code-word to feature records in 20km traversal [66]. In this situation entry and retrieval of feature vectors should be handled through optimized data access methods such as the one proposed in [61].

C. General concerns

Viewpoint invariance is one of the major concerns in place recognition. This is partially the result of “looking at a place from an external viewer’s stand point through a limited field of view”. When this approach is combined with landmark recognition, the combination remains satisfactory for most object recognition techniques. However expecting this setting to

work for place recognition introduces a logical dispute. Assume a robot is exactly located in front of a landmark and a camera with limited field-of-view, which is used as the sensor, is directed towards the landmark. Recognition of the landmark in this situation is easy, and is taken equivalent to recognition of the place. Now, if the camera faces oppositely, the mentioned landmark falls out of the field of view, and as a consequence, landmark is not recognized, and so is not the place. In both scenarios, the observation is made exactly in the same place, but the recognition results are very different. This is another reason why landmark recognition cannot necessarily be taken equivalent to place recognition. One suggestion to overcome this situation is to “look from inside of the place out, through the entire field of view”. This is achievable by using the panoramic view camera (2D) or 360 degree LiDAR (3D).

Table 1- Summary of challenges, sources, and the suggested replacement strategies

Challenge	Scope	Source	Suggested Replacement
Drift	SLAM-based place recognition	Propagation of incremental mapping error through time	Limit this approach to indoor, and short term applications.
Excessive storage and computation demands	point-wise SLAM-based place recognition	Comparison of all points of the current observation with the entire history	Limit this approach to indoor, and short term applications.
Feature selection and feature matching	feature based SLAM for place recognition	Enough number of expected type of features are not guaranteed to exist consistently in all observations.	Supplement with external referencing.
Landmark recognition	place recognition through image retrieval	Not all places contain outstanding landmarks.	Do not take landmark recognition equivalent to place recognition.
Feature selection and feature matching	place recognition through image retrieval	Enough number of expected type of features are not guaranteed to exist consistently in all observations.	Use a non-feature-based descriptor.
Few features do not fully represent places	place recognition through image retrieval	Unlike the case of object recognition, a set of features cannot fully represent a place.	Use a holistic place descriptor.
Bag-of-words dictionary grows excessively	place recognition through image retrieval	Feature vectors stored in the bag-of-words model can reach millions very quickly.	Use advanced data access models to store and retrieve data. Use non-feature-based techniques
Viewpoint invariance	general	Looking at a place from an external viewer’s stand point through a limited field of view	Look from inside of the place out, through the entire field of view.
Environmental changes	general	Color is affected by illumination along the day, and appearance changes in different seasons along the year.	Use depth information.
Depth accuracy, range, and resolution	general	Stereo vision, infrared reflection, and ultrasonic time of return measurements are not producing satisfactory performance for this application.	Use LiDAR to capture depth information
False recognition	general	Frame to frame comparison is prone to false recognition.	Use group comparison, and study places inside sequences.
General Definition for the term “place”	general	Existing place definitions do not include all types of places.	A place is the subset of the three dimensional environment in which, any partial observation made across the entire field-of-view, holds consistent descriptor that is distinctively different compared to the descriptors computed from its surrounding.

The other concern for place recognition is the depth value. Depth value is a valuable information from the place because the color information can be affected by illumination changes throughout the day, and the appearance of the place varies in different seasons along the year. Therefore, depth is the invariant information, which maintains its consistency throughout the time. Depth information can be captured through different techniques such as stereo vision, infrared reflection, ultrasonic time of flight measurement, and through LiDAR. While LiDAR can produce precise measurements in long ranges, none of the other techniques can perform the same. Therefore, it is highly recommended to use LiDAR to capture depth information for place recognition applications.

The last but not the least is the false detection issue. This problem is mainly rooted in the frame to frame comparison approach. If two places are supposed to be compared, several observations made from one place should be compared to several observations made from the other place. Group comparison reduces the chance of false detection. One other suggestion would be to study places in sequences similar to the approaches suggested by [67] and [68].

Defining the term “place” through landmarks, features, or precise coordinate location, have limitations to include all places. Taking into account all the mentioned limitations, the proposed definition of the term “place” that grounds the basis for proposition of a place recognition methodology that addresses all the mentioned concerns, should be at descriptor level by taking into account the analysis over all the observation made at different spots of the place. Therefore, the suggested definition is as follows.

“A place is the subset of the three dimensional environment in which, any partial observation made across the entire field-of-view, holds consistent descriptor that is distinctively different compared to the descriptors computed from its surrounding”.

IV. SUMMARY AND CONSLUSION

This paper studied the major challenges faced by conventional place recognition methodologies that do not rely on external referencing. For each concern, the causing factor was introduced, and the replacement treatment was suggested. This analysis is summarized in Table 1. To address these concerns, a novel methodology is suggested, which will appear in the author’s future publications.

V. FUTURE WORK

This paper addressed the shortcomings of the existing place recognition methodologies, and suggested different approaches to avoid facing them. Next step is to propose the step-by-step solution which takes into account all the mentioned concerns. This will be covered in our subsequent publications.

REFERENCES

- [1] S. Thrun, W. Burgard, and D. Fox, Probabilistic robotics: MIT Press, 2005.
- [2] S. Lowry, N. Sunderhauf, P. Newman, J. J. Leonard, D. Cox, P. Corke, et al., “Visual Place Recognition: A Survey,” IEEE Transactions on Robotics, vol. 32, pp. 1-19, 2016.
- [3] T. Song, N. Capurso, X. Cheng, J. Yu, B. Chen, and W. Zhao, “Enhancing GPS With Lane-Level Navigation to Facilitate Highway Driving,” IEEE Transactions on Vehicular Technology, vol. 66, pp. 4579-4591, 2017.
- [4] D. G. Lowe, “Object recognition from local scale-invariant features,” in Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, 1999, pp. 1150-1157 vol.2.
- [5] I. Radoi, G. Gutu, T. Rebedea, C. Neagu, and M. Popa, “Indoor Positioning inside an Office Building Using BLE,” in 2017 21st International Conference on Control Systems and Computer Science (CSCS), 2017, pp. 159-164.
- [6] Q. D. Vo and P. De, “A Survey of Fingerprint-Based Outdoor Localization,” IEEE Communications Surveys & Tutorials, vol. 18, pp. 491-506, 2016.
- [7] Y. Gu and S. Kamijo, “GNSS positioning in deep urban city with 3D map and double reflection,” in 2017 European Navigation Conference (ENC), 2017, pp. 84-90.
- [8] H. Liu, Y. Zhang, X. Su, X. Li, and N. Xu, “Mobile Localization Based on Received Signal Strength and Pearson's Correlation Coefficient,” International Journal of Distributed Sensor Networks, vol. 11, p. 157046, 2015.
- [9] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, et al., “Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age,” IEEE Transactions on Robotics, vol. 32, pp. 1309-1332, 2016.
- [10] A. C. Murillo, J. J. Guerrero, and C. Sagues, “SURF features for efficient robot localization with omnidirectional images,” in Proceedings 2007 IEEE International Conference on Robotics and Automation, 2007, pp. 3901-3907.
- [11] M. H. A. Hamid, A. H. Adom, N. A. Rahim, and M. H. F. Rahiman, “Navigation of mobile robot using Global Positioning System (GPS) and obstacle avoidance system with commanded loop daisy chaining application method,” in 2009 5th International Colloquium on Signal Processing & Its Applications, 2009, pp. 176-181.
- [12] J. Aulinas, Y. Petillot, J. Salvi, and X. L. A., “The SLAM problem: a survey,” presented at the Proceedings of the 2008 conference on Artificial Intelligence Research and Development: Proceedings of the 11th International Conference of the Catalan Association for Artificial Intelligence, 2008.
- [13] H. Durrant-Whyte and T. Bailey, “Simultaneous localization and mapping: part I,” IEEE Robotics & Automation Magazine, vol. 13, pp. 99-110, 2006.
- [14] T. Bailey and H. Durrant-Whyte, “Simultaneous localization and mapping (SLAM): part II,” IEEE Robotics & Automation Magazine, vol. 13, pp. 108-117, 2006.
- [15] KAARTA 3D Mapping. Available: <http://www.kaarta.com/>
- [16] B. M. F. d. Silva, R. S. Xavier, T. P. d. Nascimento, and L. M. G. Gonsalves, “Experimental evaluation of ROS compatible SLAM algorithms for RGB-D sensors,” in 2017 Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR), 2017, pp. 1-6.
- [17] E. Mendes, P. Koch, and S. Lacroix, “ICP-based pose-graph SLAM,” in 2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), 2016, pp. 195-200.
- [18] P. de la Puente and D. Rodriguez-Losada, “Feature based graph-SLAM in structured environments,” Autonomous Robots, vol. 37, pp. 243-260, 2014.
- [19] E. Mihankhah, H. D. Taghirad, A. Kalantari, E. Aboosaeedan, and H. Semsarilar, “Line matching localization and map building with least square,” in 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2009, pp. 1734-1739.
- [20] J. Engel, T. Schöps, and D. Cremers, “LSD-SLAM: Large-Scale Direct Monocular SLAM,” Cham, 2014, pp. 834-849.
- [21] J. Saarinen, H. Andreasson, T. Stoyanov, J. Ala-Luhtala, and A. J. Lilienthal, “Normal Distributions Transform Occupancy Maps: Application to large-scale online 3D mapping,” in 2013 IEEE International Conference on Robotics and Automation, 2013, pp. 2233-2238.

- [22] E. Takeuchi and T. Tsubouchi, "A 3-D Scan Matching using Improved 3-D Normal Distributions Transform for Mobile Robotic Mapping," in 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006, pp. 3068-3073.
- [23] D. Schleicher, L. M. Bergasa, M. Ocana, R. Barea, and E. Lopez, "Real-time hierarchical GPS aided visual SLAM on urban environments," in 2009 IEEE International Conference on Robotics and Automation, 2009, pp. 4381-4386.
- [24] R. Tron and R. Vidal, "Distributed 3-D Localization of Camera Sensor Networks From 2-D Image Measurements," *IEEE Transactions on Automatic Control*, vol. 59, pp. 3325-3340, 2014.
- [25] Z. Zhongyang, L. Yan, and J. Wang, "LiDAR point cloud registration based on improved ICP method and SIFT feature," in 2015 IEEE International Conference on Progress in Informatics and Computing (PIC), 2015, pp. 588-592.
- [26] R. Paul and P. Newman, "FAB-MAP 3D: Topological mapping with spatial and visual appearance," in 2010 IEEE International Conference on Robotics and Automation, 2010, pp. 2649-2656.
- [27] D. M. Cole and P. M. Newman, "Using laser range data for 3D SLAM in outdoor environments," in Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006., 2006, pp. 1556-1563.
- [28] H. Zhu, J. B. Weibel, and S. Lu, "Discriminative Multi-modal Feature Fusion for RGBD Indoor Scene Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2969-2976.
- [29] M. Labbé and F. Michaud, "Online global loop closure detection for large-scale multi-session graph-based SLAM," in 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2014, pp. 2661-2666.
- [30] C. Cadena, D. Galvez-López, J. D. Tardos, and J. Neira, "Robust Place Recognition With Stereo Sequences," *IEEE Transactions on Robotics*, vol. 28, pp. 871-885, 2012.
- [31] M. Hassaballah and A. I. Awad, "Detection and Description of Image Features: An Introduction," in *Image Feature Detectors and Descriptors: Foundations and Applications*, A. I. Awad and M. Hassaballah, Eds., ed Cham: Springer International Publishing, 2016, pp. 1-8.
- [32] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 679-698, 1986.
- [33] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey vision conference*, 1988, p. 10.5244.
- [34] P. R. Beaudet, "Rotationally invariant image operators," in Proceedings of the 4th International Joint Conference on Pattern Recognition, 1978, pp. 579-583.
- [35] T. Lindeberg, "Feature Detection with Automatic Scale Selection," *International Journal of Computer Vision*, vol. 30, pp. 79-116, 1998.
- [36] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, 2004.
- [37] K. Grauman and B. Leibe, "Visual Object Recognition," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 5, pp. 1-181, 2011.
- [38] K. Mikolajczyk and C. Schmid, "Scale & Affine Invariant Interest Point Detectors," *International Journal of Computer Vision*, vol. 60, pp. 63-86, 2004.
- [39] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, et al., "A Comparison of Affine Region Detectors," *International Journal of Computer Vision*, vol. 65, pp. 43-72, 2005.
- [40] A. Hubeli and M. Gross, "Multiresolution feature extraction for unstructured meshes," in Visualization, 2001. VIS '01. Proceedings, 2001, pp. 287-294.
- [41] K. Hildebrandt, K. Polthier, and M. Wardetzky, "Smooth feature lines on surface meshes," presented at the Proceedings of the third Eurographics symposium on Geometry processing, Vienna, Austria, 2005.
- [42] K. Watanabe and A. G. Belyaev, "Detection of Salient Curvature Features on Polygonal Surfaces," *Computer Graphics Forum*, vol. 20, pp. 385-392, 2001.
- [43] K. Demarsin, D. Vanderstraeten, T. Volodine, and D. Roose, "Detection of closed sharp edges in point clouds using normal estimation and graph theory," *Comput. Aided Des.*, vol. 39, pp. 276-283, 2007.
- [44] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, pp. 433-459, 2010.
- [45] J. Wang, Z. Dodds, and W. L. Miranker, "Principal Component Analysis for Place Recognition," *Journal of Neural, Parallel, and Scientific Computations*, 1996.
- [46] L. He, X. Wang, and H. Zhang, "M2DP: A novel 3D point cloud descriptor and its application in loop closure detection," in 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2016, pp. 231-237.
- [47] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," in 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., 2003, pp. II-257-II-263 vol.2.
- [48] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," in *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision*, Graz, Austria, May 7-13, 2006. Proceedings, Part I, A. Leonardis, H. Bischof, and A. Pinz, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 404-417.
- [49] J. C. Goswami and A. K. Chan, *Fundamentals of Wavelets: Theory, Algorithms, and Applications*: Wiley Publishing, 2011.
- [50] A. P. Ashbrook, N. A. Thacker, P. I. Rockett, and C. I. Brown, "Robust recognition of scaled shapes using pairwise geometric histograms," presented at the Proceedings of the 6th British conference on Machine vision (Vol. 2), Birmingham, United Kingdom, 1995.
- [51] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 509-522, 2002.
- [52] T. Cieslewski, E. Stumm, A. Gawel, M. Bosse, S. Lynen, and R. Siegwart, "Point cloud descriptors for place recognition using sparse visual information," in 2016 IEEE International Conference on Robotics and Automation (ICRA), 2016, pp. 4830-4836.
- [53] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based visual-inertial odometry using nonlinear optimization," *The International Journal of Robotics Research*, vol. 34, pp. 314-334, 2015.
- [54] M. Bosse and R. Zlot, "Place recognition using keypoint voting in large 3D lidar datasets," in 2013 IEEE International Conference on Robotics and Automation, 2013, pp. 2677-2684.
- [55] M. Bosse and R. Zlot, "Keypoint design and evaluation for place recognition in 2D lidar maps," *Robotics and Autonomous Systems*, vol. 57, pp. 1211-1224, 2009.
- [56] M. Bosse and R. Zlot, "Place Recognition Using Regional Point Descriptors for 3D Mapping," in *Field and Service Robotics: Results of the 7th International Conference*, A. Howard, K. Iagnemma, and A. Kelly, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 195-204.
- [57] T. Fiolka, J. Stückler, D. A. Klein, D. Schulz, and S. Behnke, "Distinctive 3D surface entropy features for place recognition," in 2013 European Conference on Mobile Robots, 2013, pp. 204-209.
- [58] T. Fiolka, J. Stückler, D. A. Klein, D. Schulz, and S. Behnke, "SURE: Surface Entropy for Distinctive 3D Features," in *Spatial Cognition VIII: International Conference, Spatial Cognition 2012*, Kloster Seeon, Germany, August 31 – September 3, 2012. Proceedings, C. Stachniss, K. Schill, and D. Uttal, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 74-93.
- [59] R. A. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval: Addison-Wesley Longman Publishing Co., Inc.*, 1999.
- [60] K. L. Ho and P. Newman, "Detecting Loop Closure with Scene Sequences," *International Journal of Computer Vision*, vol. 74, pp. 261-286, 2007.
- [61] D. Galvez-López and J. D. Tardos, "Bags of Binary Words for Fast Place Recognition in Image Sequences," *IEEE Transactions on Robotics*, vol. 28, pp. 1188-1197, 2012.

- [62] S. Satkin and M. Hebert, "3DNN: Viewpoint Invariant 3D Geometry Matching for Scene Understanding," in 2013 IEEE International Conference on Computer Vision, 2013, pp. 1873-1880.
- [63] A. Singhal, N. Srivastava, and R. Mishra, "Hiding Signature In Colored Image," in 2006 International Symposium on Communications and Information Technologies, 2006, pp. 446-450.
- [64] E. Mihankhah and D. Wang, "Environment characterization using Laplace eigenvalues," in 2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV), 2016, pp. 1-6.
- [65] A. C. Murillo and J. Kosecka, "Experiments in place recognition using gist panoramas," in 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, 2009, pp. 2196-2203.
- [66] G. Schindler, M. Brown, and R. Szeliski, "City-Scale Location Recognition," in 2007 IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1-7.
- [67] M. J. Milford and G. F. Wyeth, "SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights," in 2012 IEEE International Conference on Robotics and Automation, 2012, pp. 1643-1649.
- [68] K. L. Ho and P. Newman, "Loop closure detection in SLAM by combining visual and spatial appearance," *Robotics and Autonomous Systems*, vol. 54, pp. 740-749, 2006.