

A Hidden Markov Model approach for Voronoi Localization

Jie Song, Ming Liu*

Autonomous Systems Lab, ETH Zurich, Switzerland

jsong@student.ethz.ch

* Hong Kong University of Science and Technology

eelium@ust.hk

Abstract—Localization is one of the fundamental problems for mobile robots. Hence, there are several related works carried out for both metric and topological localization. In this paper, we present a lightweight technique for on-line robot topological localization in a known indoor environment. This approach is based on the Generalized Voronoi Diagram (GVD). The core task is to build local GVD to match against the global GVD using adaptive descriptors. We propose and evaluate a concise descriptor based on geometric constraints around meeting points on GVD, while adopting Hidden Markov Model (HMM) for inference. Tests on real maps extracted from typical structured environment using range sensor are presented. The results show that the robot can be efficiently localized with minor computational cost based on sparse measurements.

I. INTRODUCTION

Autonomous robots need to determine its position within its work space, answering the question “where am I?” Localization, a process that calculates the position of a mobile robot relative to its environment, is the foundation to perform other common tasks. Hence, the research on robot localization has received considerable attention.

In general, there are two main categories to represent the environment recognized by range finders: metric-based and topological approaches. As for the metric-based approach, geometric entities of the environment are represented by the exact locations with respect to a reference frame, and the localization is achieved by obtaining the accurate coordinate index of the mobile robot [1], [2], [3]. On the other hand, the topological approach represents environmental entities as a graphical model composed by nodes and edges [4], [18], [6], [7]. The nodes are usually extracted from places with salient features, while edges among the nodes determine the connectivity. For vision based approaches, the localization is mostly achieved by feature matching and geometrical triangulation [19]. In this proposed approach, the localization task aims at finding the topological region [8], [23] to which the current robot pose belongs. It is closely related to topological decomposition problem [24]. This goal may sound vague, however it is good enough for local environment modeling tasks e.g. local pointcloud based semantic mapping [22] or local data retrieval [26], [25].

Furthermore, comparing with explicit metric localization, the proposed scheme requires lower computational cost, and is easier to recover for map lost. The occupancy grid map [9], feature-based methods [10] and appearance-based maps [20] are typical examples of explicit metrical localization approaches. However, to implement the precise localization,

the autonomous robot has to handle large amount of redundant information. Contrarily, due to the compact and abstracted scheme, a topological localization approach can massage the required data more easily as we previously discussed in [21].

This work addresses topological localization by using GVD[4]. An environment model is first created, which involves partitioning the Voronoi graph into several topological subregions. Then the robot can try to localize by comparing the local Voronoi diagram with the global one. The Voronoi diagram is directly calculated based on the grid map obtained from range sensors. At the same time, the topological nodes are defined according to the structure of the Voronoi diagram.

The rest of this paper is structured as follows. In the section II, state of the art will be discussed and compared. In section III, we propose a topological segmentation method based on global GVG map. Then geometrical information based descriptor for GVG in section IV. It is evaluated and parametrized by sampler GVG in section V. In section VI, a Hidden Markov Model is proposed for the inference of the model, followed by experiments for the proposed method and discussions. Finally, we conclude in section VI.

II. RELATED WORK

A. Topological localization

Due to the advantage of huge reduction in both the space and time complexity, topological localization is widely studied. They can be characterized by different methods of feature extraction. One popular example is based on a sonar grid map [8]. In this method, localization is performed by comparing the current local grid map with the grid map constructed during the modeling procedure. Though topological model can be quickly extracted from a grid map by using cell decomposition and normalized graph cut, there is a big limit that the region segmentation is rather unstable. This is because spectral clustering was adopted for topological region segmentation. Even predefined the number of regions to divide, results can vary occasionally. In our previous report, the conducted problems were detailed analysed [23]. Brunskill et al proposed a localization based on boosting over a set of local features extracted from raw laser scan [11]. They create the topological map based on second largest eigenvalue of spectral analysis, which lead to large and non-organized topological segmentation. Blanco et al used a naive local pattern-matching algorithm for GVG matching [12], which is infeasible to adapt to rotation. Another Voronoi localization algorithm is to use stochastic mapping [13]. Comparing with that, we use HMM

in this paper to refine the matching sequence. A widely cited method was proposed by Choset et al [4], [14]. However, the global refinement was not dealt with. In this work, we use Viterbi algorithm to refine the final output.

Other less related approaches are to use multi sensor features, e.g. extracted from the images obtained via an omnidirectional camera. Mostly, two types of feature namely global features extracted from whole image and local features obtained from subregions are considered [15]. This concept is adopted in GVG descriptor extraction in our approach.

B. Voronoi Diagram and further application

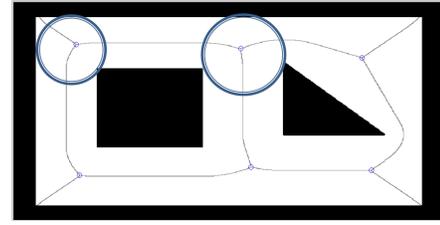
The Voronoi diagram (VD) and GVD are both geometric structures formed by sets of points in R^n (in my application R^2). It consists of a set of points that have at least two minimally distant sites in terms of Euclidean distance. All points which are not part of the VD are closest to exactly one particular point obstacle. It is composed of line segments and half-lines, which are both formed of points with exactly two equal minimal distance to the point obstacles. These lines meet at so-called meet points, which have at least three equal minimal distance to the point obstacles. The GVD extends the idea of the VD from point obstacles to a set of arbitrary geometric objects (e.g., lines, arcs). There are many applications based on the VD and GVD. Since the definition of VD guarantees that robots can stay safe along the edges, some people use it to perform robot path planning and navigation. A typical example is to use Hidden Markov Models for structural information extraction at distinct indoor places, like corridor, hall and intersection [16].

III. TOPOLOGICAL MODELING

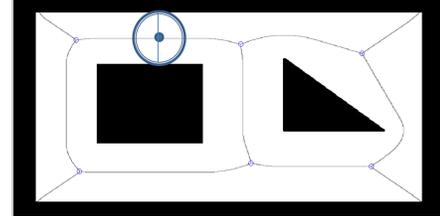
GVD is a collection of all the equally distant points to neighbouring obstacles. In other words, they are the centers of circles which are tangent to surrounding obstacles. As for a generalized description, the obstacles can be either points or any other arbitrary geometric shapes. As for the points whose circles are tangent to at least three distinct obstacles, we call them meet points of GVD. The union of other points on the GVD are defined as edges. Different meet points are connected by explicit edges, while different edges intersect at meet points. The following steps are implement to define the topological regions for a given map.

- 1) Extract the GVD for the given map;
- 2) For all pairs of adjacent meet points, tracking along the edge that connects them;
- 3) Locate the half-way point on this edge and draw the circle which is tangent to the surrounding obstacles;
- 4) Connect the two closest obstacle points and connect them as the separation boarder of two regions;
- 5) Go to 1), till all regions are extracted.

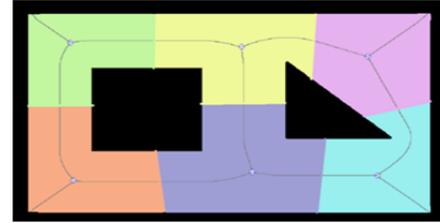
Figure 1 illustrates the process to divide a target map into distinct topological regions represented by different colors. Since the meet points usually appear in the corners or intersections, it is an intuitive way to define the topological regions according to the location of meet points.



(a) Find an arbitrary pair of two neighboring meet points



(b) The half-way point on the edge



(c) Topological regions by connecting base points on the obstacles

Fig. 1. Region defining: figure a is the GVD of a structured map. First find any two adjacent meet points. Then the middle point on this edge is extracted. Finally all the base points on the obstacles are found out and connected to form the GVD

IV. DESCRIPTOR BUILDING AND MATCHING ALGORITHMS

The critical procedures for localization are to build the descriptors for meet points and to find proper matching algorithms. A direct approach is to record the relative coordinates of the edge points extended from one meet point. Then the pattern matching technique can be performed [12], originated from image processing techniques [17]. However, to compare the local pattern with the global templates, the Local Voronoi Diagram (LVD) map should be aligned with the global map according to the current pose of the robot. An orientation-invariant geometrical characteristics is still required.

Several features can be extracted from meet points of GVD. We choose the following four major features. The most typical one is number of edges emanating from a meet point. Second is the distance to the closest obstacles at the meet point. The third one is to calculate the relative departure angles of GVD edges extended from the meet points. The last is the mean curvatures for the edges to certain distance (parameterized in figure 8).

In order to match two meet points, e.g. a meet point from LVD with the global ones, the following steps are performed.

A. Valid the number of edges emanating from the meet point

As shown in figure 2, the number of edges is the most direct and intuitive characteristic. This means the sensor noise and

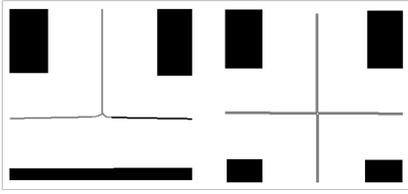


Fig. 2. The most distinguishable feature is the number of edges emanating from a meet point. The left one has three while the right one has four edges.

position uncertainty will not affect this feature much. If the number of edges of the local meet point is different from the corresponding number of the global one, we can say that they cannot be matched. If the numbers are the same, the following steps are carried out.

B. Distance to the closest obstacles at the meet point

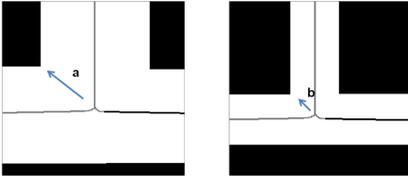


Fig. 3. Even though the geometric shape of the edges emanating from these two meet points are quite similar, the distances to the closest obstacles are different enough to distinguish them.

As shown in figure 3, even though their geometric shapes are similar, the shortest distance to the obstacle for the first meet point (a) is larger than the second one's (b). This distance measurement is used to quickly eliminate negative candidate meet points. Hence, we can set a threshold level according to the reliability of the sensor we use. If the value $|\frac{a-b}{b}|$ is greater than an empirical threshold, we can conclude that they are from distinct places. On the other hand, if the value is within this range, we should continue to execute next validations.

C. Angular features and mean curvature

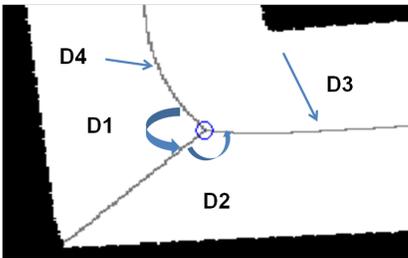


Fig. 4. Descriptor vector is consist of four dimensions: largest angle between edges, smallest angle, largest mean curvature of the edges, smallest mean curvature.

As shown in figure 4, we construct a descriptor vector that consists of four dimensions. The first two dimensions (minimum angle and maximum angle) represent the angle feature of the edges emanating from the meet point. The other two dimensions (minimum curvature and maximum curvature) are the mean curvature information. Several typical criteria are

to be adopted in order to compare two descriptors. In order to find out which one is most appropriate for this application, twenty meet points with their edges were arbitrarily generated and the descriptor vectors for all of them were calculated. Then we cross match these sampled descriptor vectors using different distance criteria.

$$\theta = \arccos \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \right).$$

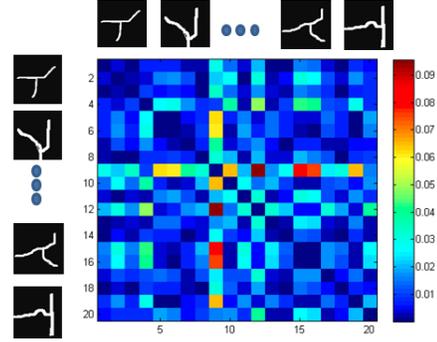


Fig. 5. Evaluation of cross matching by angle between two vectors.

$$d(\mathbf{a}, \mathbf{b}) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2}$$

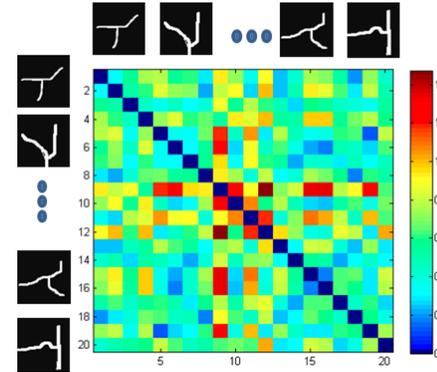


Fig. 6. Evaluation of cross matching by Euclidean distance

1) *Separation angle by inner-product*: The figure 5 shows the result of cross comparison using the separation angles depicted at the top figure. The color on the diagonal is dark blue while the colors on two sides are similar. This means the angles between the descriptor vectors are so small that it is difficult to distinguish them in this way.

2) *Euclidean distance*: The figure 6 is the evaluation result by using Euclidean distance. The values on the two sides are more distinct from the values on the diagonal, which implies it is more reliable for matching.

3) *Mahalanobis distance*: Since the evaluation is offline, and all descriptors are available for evaluation. For the purpose of generalization, Mahalanobis distance is adopted, which introduces the scaling factor for each dimension. The result is shown in figure 7. We could see that Mahalanobis distance can better distinguish the differences.

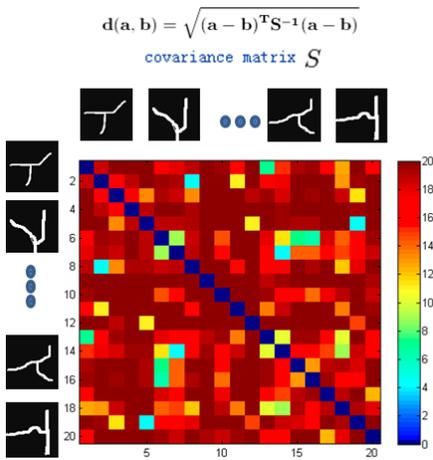


Fig. 7. Evaluation of cross matching by Mahalanobis distance

V. PARAMETERIZATION

Using these offline results, we could train the parameters that are used for real experiment. Especially the calculation of mean curvatures highly depends on the tracking distance apart from the meet points. In other words, it is the area in which we calculate the descriptor vector for the target meet point. This evaluation can be implemented by a recall-precision test.

Definitions of recall and precision are shown in equation 1. First, fifty local descriptors for one target global meet point were built by using virtual laser sensor fifty times. Then another fifty local descriptors for another global meet point were calculated in the same way. The threshold value for matching can be set as the Mahalanobis distance between those two original global descriptor vectors.

Figure 8 shows the recall-precision result of these one hundred descriptor vectors to the target vector. We can see that the matching recall decreases while the matching precision goes up with the increasing of track distance. It means, if the robot just focuses on a small range of area, the meet points have high chance to be mismatched. On the other hand, if the features are extracted in a bigger range, even though the precision for matching is high, the value of matching recall drops dramatically. So the optimal track distance should be the one maximizing the multiplication of matching recall value and precision value. This optimized track distance is just for the description of this meet point. The repetitive work has to be done for all the meet points in the given global map. The resulting track distance is the mean of all the optimal track distances.

$$\begin{aligned} \text{precision} &= \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \\ \text{recall} &= \frac{\text{true positive}}{\text{true positive} + \text{false negative}} \end{aligned} \quad (1)$$

VI. SEQUENTIAL LOCALIZATION BY HIDDEN MARKOV MODEL

Since the state space for the topological map is discrete and neighborhood relevant, a Hidden Markov Model can be

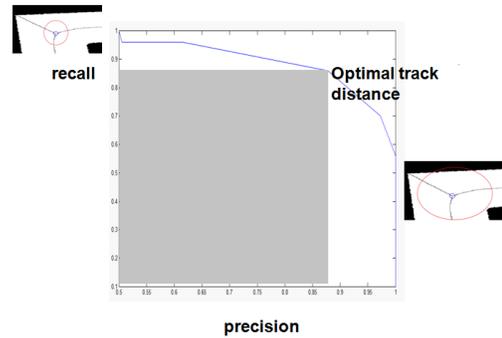


Fig. 8. Through the recall-precision test, we find out the optimal track distance to build the descriptor for this target meet point.

constructed to perform a reliable localization. The following variables are defined accordingly.

- **The state space S :**
The state space is composed of topological regions characterized by the meet points. We denote the number of states in the model by N .
- **The initial distribution π_i :**
If there is no prior information for the location of the robot, the initial distribution among S is uniform.
- **The state transition probability matrix $A = \{a_{ij}\}$:**
On the condition that there is no trained preference for the robot, the robot is supposed to have equalized probability to stay in the previous region and to move into its neighboring topological regions. This is demonstrated in the figure 9.

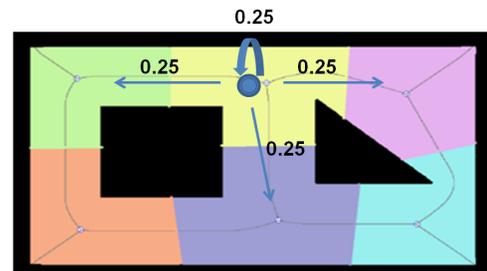


Fig. 9. transition matrix: the same probability to stay in the same region or to go to the neighboring regions.

If the robot has predefined preference, for example, it may high chance to turn left when it is located at special intersection. In this case, the transition matrix should be trained using Baum-Welch algorithm.

- **Observation sequence:**
It is denoted by $O = \{O_1, O_2, \dots, O_t\}$. It is comprised of a sequence of descriptors extracted from local laser observations (see: LVD in [12]).
- **The observation matrix $B = \{b_j(k)\}$:**
Since the observation space is a continuous space(k represents the observation), the observation matrix is supposed to be computed for each time step. As discussed above, the local map can be obtained by simulating a laser range sensor. Then local GVD is calculated and the

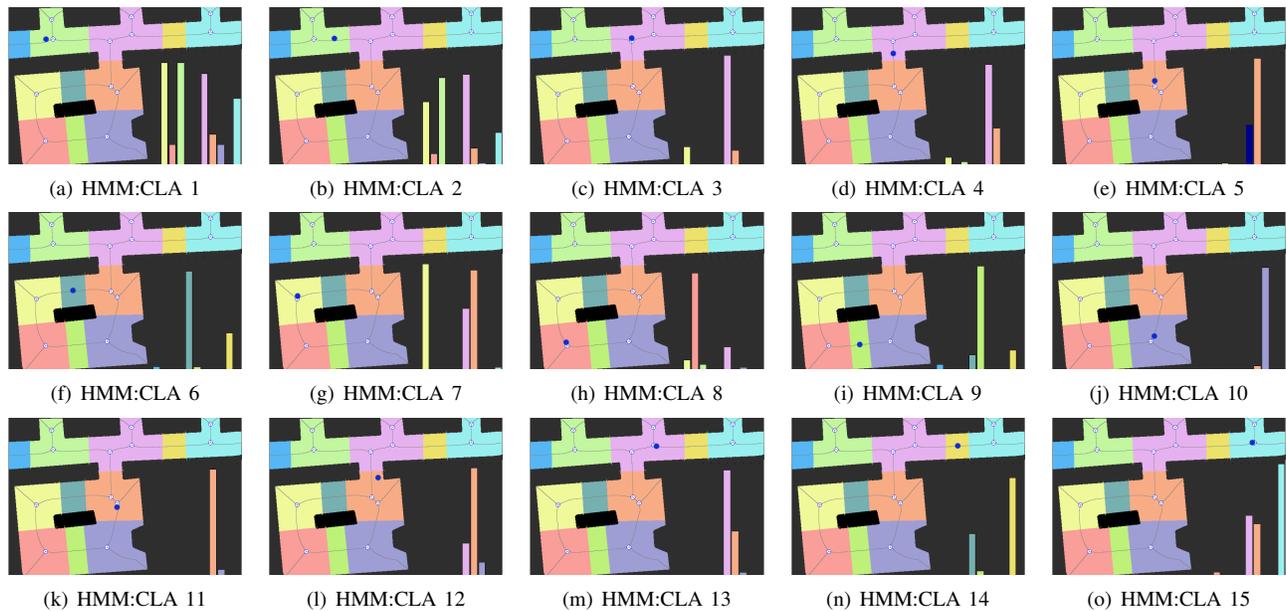


Fig. 11. HMM simulation process for CLA building: the blue dot represents the position of the robot. The bar height represents the probability staying in the corresponding region for the robot, that is: $P(S^t|O^0, \dots, O^t)$. At the beginning, the robot was quite confused by some similar global meet points, but after it saw more meet points it corrected the mistake based on both observation model and state transition matrix.

- [2] F. Lu and E. Miliotis, "Robot pose estimation in unknown environments by matching 2d range scans," *Journal of intelligent & robotic systems*, vol. 18, no. 3, pp. 249–275, 1997.
- [3] S. Thrun, W. Burgard, and D. Fox, "A probabilistic approach to concurrent mapping and localization for mobile robots," *Autonomous Robots*, vol. 5, no. 3, pp. 253–271, 1998.
- [4] H. Choset and J. Burdick, "Sensor-based exploration: The hierarchical generalized voronoi graph," *The International Journal of Robotics Research*, vol. 19, no. 2, pp. 96–125, 2000.
- [5] B. Kuipers and Y. Byun, "A robust, qualitative method for robot spatial learning," in *Proceedings of the AAAI*, vol. 88, 1988, pp. 774–779.
- [6] M. Liu, F. Colas, F. Pomerleau, and R. Siegwart, "A Markov semi-supervised clustering approach and its application in topological map extraction," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012. (IROS 2012)*, 2012.
- [7] M. Liu, C. Pradalier, F. Pomerleau, and R. Siegwart, "The role of homing in visual topological navigation," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012. (IROS 2012)*, 2012.
- [8] J. Choi, M. Choi, S. Nam, and W. Chung, "Autonomous topological modeling of a home environment and topological localization using a sonar grid map," *Autonomous Robots*, vol. 30, no. 4, pp. 351–368, 2011.
- [9] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, 1989.
- [10] J. Leonard and H. Durrant-Whyte, "Simultaneous map building and localization for an autonomous mobile robot," in *Intelligent Robots and Systems '91: Intelligence for Mechanical Systems, Proceedings IROS '91. IEEE/RSJ International Workshop on*, IEEE, 1991, pp. 1442–1447.
- [11] E. Brunskill, T. Kollar, and N. Roy, "Topological Mapping Using Spectral Clustering and Classification," in *International Conference on Intelligent Robots and Systems*. IEEE, 2007, pp. 3491–3496.
- [12] D. Blanco, B. Boada, and L. Moreno, "Localization by voronoi diagrams correlation," in *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, vol. 4. IEEE, 2001, pp. 4232–4237.
- [13] J. Wallgrün, *Hierarchical voronoi graphs: Spatial representation and reasoning for mobile robots*. Springer Verlag, 2010.
- [14] H. Choset and K. Nagatani, "Topological simultaneous localization and mapping (slam): toward exact localization without explicit localization," *Robotics and Automation, IEEE Transactions on*, vol. 17, no. 2, pp. 125–137, 2001.
- [15] H. Andreasson, A. Treptow, and T. Duckett, "Localization for mobile robots using panoramic vision, local features and particle filter," in *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*. IEEE, 2005, pp. 3348–3353.
- [16] B. Boada, D. Palazon, D. Blanco, and L. Moreno, "Voronoi based place recognition using hidden markov models," *Proc. 15th IFAC World Congr. Int. Fed. Autom. Control*, p. 2002, 2002.
- [17] J. Parker, *Practical computer vision using C*. John Wiley & Sons, Inc., 1993.
- [18] MING LIU, CEDRIC PRADALIER, ROLAND SIEGWART: *Visual Homing from Scale with an Uncalibrated Omnidirectional Camera*. IEEE Transactions on Robotics, 2013. DOI: 10.1109/TRO.2013.2272251
- [19] MING LIU, CEDRIC PRADALIER, FRANCOIS POMERLEAU, ROLAND SIEGWART: *Scale-only Visual Homing from an Omnidirectional Camera*. IEEE International Conference on Robotics and Automation (ICRA), 2012.
- [20] MING LIU, ROLAND SIEGWART: *Topological mapping and scene recognition with lightweight color descriptors for omnidirectional camera*. IEEE Transactions on Robotics, 2013. DOI: 10.1109/TRO.2013.2272250
- [21] MING LIU, DAVIDE SCARAMUZZA, CEDRIC PRADALIER, ROLAND SIEGWART AND QIJUN CHEN: *Scene Recognition with Omnidirectional Vision for Topological Map using Lightweight Adaptive Descriptors*, IEEE Transactions on Robotics, 2013.
- [22] MING LIU, ROLAND SIEGWART *Information Theory based Validation for Point-cloud Segmentation aided by Tensor Voting*. IEEE International Conference on Information and Automation (ICIA), 2013
- [23] MING LIU, FRANCIS COLAS AND ROLAND SIEGWART: *Regional topological segmentation based on Mutual Information Graphs*. IEEE International Conference on Robotics and Automation (ICRA), 2011
- [24] MING LIU, FRANCIS COLAS, FRANCOIS POMERLEAU, ROLAND SIEGWART: *A Markov Semi-Supervised Clustering Approach and Its Application in Topological Map Extraction*. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2012
- [25] LUJIA WANG, MING LIU, MENG, M.Q.-H., SIEGWART, R.: Towards real-time multi-sensor information retrieval in Cloud Robotic System, Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on , vol., no., pp.21,26, 13-15 Sept. 2012 doi: 10.1109/MFI.2012.6343054
- [26] LUJIA WANG, MING LIU, MENG, M.Q.-H.: Towards cloud robotic system: A case study of online co-localization for fair resource competence, Robotics and Biomimetics (ROBIO), 2012 IEEE International Conference on , vol., no., pp.2132,2137, 11-14 Dec. 2012 doi: 10.1109/ROBIO.2012.6491284