Unsupervised and Online Place Recognition for Mobile Robot based on Local Features

Tang S.H. and G. Hamami
Universiti Putra Malaysia, Malaysia
Email: saihong@eng.upm.edu.my, ghazali.hamami@gmail.com

B. Karasfi
Universiti Putra Malaysia & Islamic Azad University, Qazvin Branch
Email: karasfi@qiau.ac.ir

D. Nakhacenia
University of Ottawa, Canada
Email: dania@uottawa.ca

Abstract—Place recognition approaches have been used for solving topological mapping and localization problems. These approaches are usually performed in supervised and offline mode. In this paper, a robust appearance-based unsupervised and online place recognition algorithm, which is inspired from online sequential clustering methods, is introduced. This method combines several image features using Speedup Robust Features (SURF) by accumulating them into a union form of features inside each place cluster. In this method, the mobile robot captures omnidirectional visual information and converts them into topological place clusters. Experimental results show the robustness, accuracy, and efficiency of the method as well as its ability to create topological place clusters for solving mapping and qualitative localization problems. The performance of the developed system is remarkable in term of recognition precision performance.

Index Terms—place recognition, SURF features, online clustering, environment modeling, and topological localization.

I. INTRODUCTION

The effect of urban and industrial life is using automatic machine such as mobile robots for help and assistance to the people for doing their jobs smarter and faster. Mobile robots have various applications in different fields, such as services at home, industrial manufacturing in industry, military, transportation, healthcare in hospitals, robotic fire fighters, surveillance robots to exploration of the deep sea, planets and space applications and so on [1]-[4]. Different applications mean that the mobile robots must be able to work in various environments. The working environments could be categorized according to the robot user side point of view. Environmental representation is a basic problem in mobile robot navigation system. In the past, several researches have been established to build an accurate and complete metric or topological map of the environment based on the data gathered by the mobile robot [3], [5]-[8].

Mobile robots needs to move to perform their tasks, so having a robust navigation system is essential. "Where am I?" is a fundamental question in the navigation system. Answering to this question is a research interest, which is related to mapping and localization methods. The way of mapping and localized the robot inside a map depends on various environmental and robot platform’s parameters. For example, the size of working environment. This research focuses on large-scale environment, which is related to the global localization and mapping approaches.

Mapping and localization methods symbolized the mobile robot environments by different ways. The mobile robot perception based on visually information's of current place, which is named place recognition, is used for topological mapping and localization methods [2], [9]-[11]. For a kind of assistive mobile robots that are designed to cooperate and communicate with their users, it is more desirable to have a similar localization and mapping system with the human perception. Regarding to achieve to this target appearance or visualization of places could be used for creating a topological graph of places. This graph could be utilized for qualitatively localized the mobile robot place. This map shows the spatial relationship between different places in the environment similar to the concept of topological representation [12]-[14].

Mapping and localization approaches required to categories’ sensory information for recognizing the mobile robot place. Place classification has been performed by a supervised learning approach to label different locations. These place classification or recognition approaches have been developed based on classification methods such as
the support vector machine (SVM) [15]-[19]. However, supervised learning methods are increase human supervision in qualitative localization and mapping process (place classification). Moreover, when the mapping and localization, methods performed based on supervised learning approaches, usually, its need to gathering the sensory information and then divide these data under human supervision, finally, feed this separated information to the learning system for creating the model of the working environment [7], [9], [17], [18]. However, if the environment has changed and/or expanded, the learning phase has to be repeated. This problem is so important when the mobile robot runs in a dynamic environment with human activities.

In this research unsupervised and online place clustering method will be introduced for localization based on place appearance. This method could be utilized to create topological map and qualitative localization process. The reference or representative image database, which is represents a place cluster-kernels, could be obtained based on common local features in a sequence of images. As an alternative to the classification (supervised learning) methods which are regularly used in place recognition [16], [17], [19], this research proposes new place clustering and recognition method based on online sequential clustering method [20] which requires less human supervision.

This article is organized as follows. Section 2 presents a short overview of the related work. Section 3 describes the background of the methods which are applicable in place clustering and recognition. In Section 4, we present our proposed method. In section 5, the experimental results are shown, and finally, we draw the conclusion in Section 6.

II. RELATED WORKS

The mobile robots which are cooperating with the humans need to have the capability of recognizing their places to run their tasks. Place recognition is a basic and fundamental requirement when the mapping and localization methods want to building the topological map of an environment. The topology of a working environment could be discovered by place recognition. Place recognition could be utilized to create the topological map of the area and assistive mobile robots use this map to find out when the robot has arrived at the requested target [16], [17], [19], [21]-[23].

Camera system as a visual sensory source has significant advantages compared to proximity sensors [24]. Low-price high-quality cameras could provide noteworthy information that is applicable for place recognition applications [3]. A survey about existing techniques with focus on vision-based mapping and localization methods in mobile robot navigation systems are accessible in [3].

Appearance based mapping and localization methods divided into two main category which are object-based and region-based place recognition [3]. Object-based place recognition systems [5], [21], [25]-[27] try to identify a place by detecting the objects belonging to that place. Region or scene-based place recognition systems [7], [9], [13], [16], [17], [19], [28]-[31] characterized the environment based on set of visual features, for example, local or global invariant features [32]-[34] extracted from the images of the environment.

In this research, region-based place recognition methods will be introduced for localization based on place appearance. Mapping and localization approaches required to categories' sensory information for recognizing the mobile robot place. Place classification of indoor environment has been performed by a supervised learning approach to label different locations [35]. A three-layer neural network trained with the back-propagation algorithm has been done for scene classification [30]. Vasudevan and Siegwart introduced a Bayesian space conceptualization and place classification method for mobile robot mapping, which are based on learning from exemplars, clustering and the use of Bayesian network classifiers [25]. Moreover, the place classification or recognition approaches which are applicable in mapping and localization algorithms have been developed based on classification methods such as the support vector machine (SVM) [15]-[19].

An unsupervised place clustering method for creating a topological mapping and qualitative localization was introduced in [31]. However, this work is still running in offline mode, and the learning and testing phase is separated. This method inspire from the k-mean clustering method and build a reference image database, which is more similar to current robots query image based on their local features. As an alternative to the offline clustering (unsupervised learning) method, and inspire from Basic sequential clustering method [20], this research proposes a new online sequential place clustering method based on the difference of the Speed-Up Robust Features (SURF) [34] local features in a sequential stream of the mobile robot environmental images.

III. BACKGROUND

In this section, we will describe the methods which are applicable to unsupervised place classification process such as image similarity measurement, Union of local Features operation and the K-adjacent-union place clustering algorithm.

A. Image Similarity Measurement

Image similarity is an important issue in place recognition methods. Calculating the distance between two images could be found out based on the distinctive features from the image scene. Local image features such as SURF could be utilized in measuring the image similarity [16]. Fig. 1(a) and Fig. 1(c) are shown an omnidirectional and unwrapped image which is captured while a mobile robot running in working environment. The SURF local features could convert an omnidirectional image to a two dimension numerical matrix which is shown in Fig. 1(b). In this matrix, each column represents a vector of numbers that describe a local feature and the column numbers shows the number of extracted local features.
Figure 1. (a) An omnidirectional image captured during the mobile robot movement, (b) Local features matrix extracted from the image, each column represents a vector which is describe a local feature, (c) Unwrapped an omnidirectional image.

The SURF-based similarity measurement is shown in Eq.1 and Eq.2.

\[ \text{Sim}(I_h, I_k) = 1 - \frac{1}{\text{dim}} \sum_{f=1}^{\text{dim}} \min_{f_m=1,...,na} \{\text{Dist}(f_h, f_k(m))\} \] (1)

where \( I_h \) and \( I_k \) are local feature sets and \( f_h \) and \( f_k \) are two single SURF features. The local feature similarity kernel could be selected from any kind of kernels. The Euclidean distance kernel has been selected over the SURF local feature that is shown in Eq.2:

\[ \text{Dist}(f_h, f_k(m)) = \sqrt{\sum_{m=1}^{F_{\text{DIM}}} (f(h,m) - f(k,m))^2} \] (2)

where \( F_{\text{DIM}} \) is a local feature's dimension which is equal to 64 for the SURF features [34]. In addition, we need to know the index of local features that correspond together. Eq.3 shows that the correspondent features can be calculated between the two sets of local features.

\[ \text{Sim}_{\text{match}}(I_h, I_k) = \arg \min_{j=1,\ldots,D_k} \left\{ \text{Sim}(I_h, I_k) \right\} \] (3)

B. Union of Local Features

Place clustering and recognition methods need to build a cluster-kernel for modeling the place clusters. Features Mean calculation process causes to lose the original value of local features; as a result, the amount of error in the next matching calculation process is increased. The union calculation is an alternative to build the cluster-kernels and the algorithm which is represent the union calculation process are shown as follow[31] :

\[
\text{Union}_{\text{Of}} \text{Cluster}(m_i, k) = \begin{cases} 
\text{Sim}_{\text{match}}(f_{i}, f_{k}) & \text{if } m_i \text{ is the cluster kernel} \\
\text{Sim}(f_{i}, f_{k}) & \text{otherwise}
\end{cases}
\]

In this algorithm \( m_i \) represents the cluster kernel \( C_i \) and \( f_i \) related to the SURF features of images \( I \) which are belonged to cluster \( C_i \). The \( w_{i} \) contains the weights value of local features \( m_i \). The \( SM \) is a two dimension vector which contains the indexes of similar features between cluster-kernel \( m_i \) and image \( I_k \). First row of \( SM \) is related to kernel cluster \( m_i \) and the second row of \( SM \) is related to image features \( I_k \). In the beginning of calculation the first features set \( I \), which is belongs to cluster \( C_i \), is considered as cluster-kernel. It then continues with all \( I_k \) that were included in cluster \( C_i \). By utilizing the \( \text{Sim}_{\text{match}} \) function the correspondence feature indexes between cluster kernel \( m_i \) and feature set \( I_k \) are extracted. These indexes are related to the repeated features, therefore the weights value of these features are increased. The unmatched features from feature set \( I_k \) are also added to cluster kernel \( m_i \) but the weights value for these new features are assigned to 1(1one). These process are repeated for all features sets \( I_k \) of cluster \( C_i \). Finally, those features with minimum weight are deleted from \( m_i \) because of lowest repetition rates in cluster \( C_i \).

C. K-adjacent-union Place Clustering Algorithm (Offline Method)

This method was introduced in [36]. It is an unsupervised and offline clustering method that clustered a sequence of images captured by mobile robot. This method inspired from the standard k-mean clustering approach. It has two different steps, which are learning and recognition. Moreover, this method used the union operation as an alternative of mean to keep the original local features in each cluster-kernel and increase the precision performance of place recognition. In addition, the authors limit the searching place cluster to the neighboring area according to the adjacent characteristic of the working environment in topological map graph. Finally, in initial step, they used the local minima points in the neighboring image similarity diagram to estimate the number of places. This method is shown as follows:

1. Apply scene change detection method over the sequence of robot environmental images and extract the number of clusters \( k \).
2. Initialize a \( k \)-partition during the sequence of images and calculate the local features union of each cluster: \( m_1, m_2, \ldots, m_k \) according to Union(Of,Cluster set \( m_i \in C_i \), for \( i = 1, \ldots, k \).
3. Compare all cluster-kernels \( M_i \) to find the similarity between them by using Eq.1.
4. According to the similarity that is extracted in step 3, join the clusters that are similar and create the topological graph of cluster-kernels which we called it as place graph.
5. Assign current cluster \( C_i \) to 1 based on prior knowledge.
6. Assign Image \( I_k \) in the dataset to the nearest neighboring cluster \( C_i \) in the place graph, i.e., \( I_k \in C_i \) when \( d(I_k, C_i) = \min_{j=1,\ldots,k} \{d(I_k, C_j)\} \).
7. Recalculate the local features union of each cluster: \( m_1, m_2, \ldots, m_k \) according to Union(Of,Cluster set \( m_i \in C_i \), for \( i = 1, \ldots, k \).
8. Repeat steps 5 to 7 until there is no change happened for each cluster by comparing the current and previous result.

The result of this method is a topological graphs that shown the relation between the place clusters. Each node in this graph stored a union of local features, which are common in different scenes that are belonged to a correspondent cluster.
IV. ONLINE SEQUENTIAL PLACE CLUSTERING (PROPOSED METHOD)

In this section we introduced our proposed method which is inspired from online sequential clustering method [20]. This method utilizes unwrapped omnidirectional images as an input data. The distance between input images and cluster-kernels calculated by the number of SURF features which are matched between input image and each of cluster-kernels (union of common local features) according to Eq.4.

\[ distance(i) = \frac{2 \cdot \text{size(sim_match(C_i,input-image))}}{N_C_i + N_{\text{input-image}}} \]  

(4)

Finding the match features performed by applying Eq.3, \( N_{C_i} \) is the number of SURF local features in cluster \( C_i \) and \( N_{\text{input-image}} \) is the number of local features in input-image. The distance\((i)\) calculated for all current cluster neighbors in topological graph. The proposed method represent as follow:

Which in this algorithm, \( T_{\text{New\_Cluster}} \) is a threshold for creating the new cluster and \( T_{\text{Update\_Cluster}} \) is a threshold for updating an existing cluster-kernel. Note that \( T_{\text{New\_Cluster}} \) is less than one and greater than \( T_{\text{Update\_Cluster}} \).

V. EXPERIMENTAL RESULTS

To test the performance of our proposed method, the COLD dataset has been selected. This dataset is available in http://cogvis.nada.kth.se/COLD. This dataset includes set of experiments, which are captured in the indoor laboratory environments in three universities with different rooms and various functionalities. Moreover, three robot platforms have been used, which are the ActivMedia-Pioneer3 in Freiburg University, the iRobot-ATRV-Mini in Ljubljana University and the ActiveMedia-PeopleBot in Saarbrücken University. All of these robot platforms have a Stand-alone and Omnidirectional camera. In this research, the unwrapped omnidirectional image has been use to perform the place clustering experiments, and the image size in all experiments is 320*240 pixels.

The accuracy of the proposed method has been investigated by comparing it with the k-mean, k-union-adjacent place and the Weighted Pair Group Method Average in Agglomerative Hierarchical (WPGMA-AH) Clustering [20] methods.

Fig. 2 represents the place recognition performances in different experiments and various working environment. The averages of clustering precision performance for k-mean, WPGMA-AH, k-union-adjacent and the proposed method are equaled to 71.05%, 78.56%, 84.17% and 94.573%. With regards to the clustering precision performance, one can appreciate the fact that the performance of the proposed method is higher than the other methods.

VI. CONCLUSION

Localization and mapping are essential elements of the navigation system which is necessary for autonomous mobile robot. Place recognition could be used in appearance based topological mapping and localization methods. Appearance-based solutions are portable and cost effective. Place recognition methods usually perform by classification methods in offline mode. In this paper, an unsupervised online place recognition method has been introduced that can utilize the omnidirectional vision system and focus on the image similarity based on the SURF local features extracted from the robot environmental images. This method can automatically cluster the visual information, and also by using the sequential adjacent characteristic of mobile agent workspace, the topological graph of the place clusters can be created. Based on experimental results over the COLD dataset, the average of recognition precision is more than 94% in different trials. The results obtained from the experimental result based on the mobile robot omnidirectional images proved that this place recognition method is robust, accurate, and applicable to the different mobile agent qualitative localization task assignments.

ACKNOWLEDGMENT

This work was supported By Universiti Putra Malaysia.

REFERENCES


BIographies

Tang Sai Hong received his PhD and BEng from Dublin City University and Universiti Pertanian Malaysia, respectively. He is an Associate Professor and teaches with the Department of Mechanical & Manufacturing Engineering, Universiti Putra Malaysia since 1997. Currently, he works in the fields of robotics, operations research and artificial intelligence.

Babak Karasfi received his Bachelor and Master degrees from Iranian Universities. Currently, he is pursuing his PhD in Universiti Putra Malaysia. He is also an academic in the Islamic Azad University, Qazvin Branch. At this moment, he is focusing in robotic motion planning and artificial intelligence.

Daniel Nakhaeinia received his Bachelor and Master degrees from Iranian University and Universiti Putra Malaysia, respectively. Currently, he is pursuing his PhD in University of Ottawa. His research focus includes robotic motion planning and artificial intelligence.

Ghazali Hamami received his Bachelor degree from Japanese University. He is a Master degree candidate of Universiti Putra Malaysia. He is also attaches with Universiti Teknologi Mara. His research focus includes robotic motion planning and artificial intelligence.