An Image-based Outdoor Place Recognition and Information Retrieval System

Hanlin Goh, Yiqun Li, Joo-Hwee Lim
Computer Vision and Image Understanding Department
Institute for Infocomm Research
21 Heng Mui Keng Terrace, Singapore 119613
{hlgoh, yqli, joohwee}@i2r.a-star.edu.sg

ABSTRACT
In image-based place recognition, an image is used to deduce the location of the viewer during image acquisition. The identified place can subsequently be used to provide additional information to the user. This demonstration paper briefly describes a system that performs image-based place recognition on outdoor images, made possible by viewer-centric data sampling and local feature-based scene identification. It also explains our proposed demonstration that displays the recognized place on a map and provides information of amenities in its vicinity.

Categories and Subject Descriptors
I.4.8 [**Image Processing and Computer Vision**]: Scene Analysis—*object recognition*; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

General Terms
Algorithms, Design

Keywords
Place Recognition, Scene Identification, Information Retrieval.

1. INTRODUCTION
Inspired by the growing popularity of digital imaging, image-based applications involving the recognition and classification of images have emerged. One such application is that of outdoor place recognition, which uses one or more acquired images of an outdoor scene to deduce where the images were acquired. In this case, the output of the system is neither a precise location with GPS coordinates, nor a broad categorization of places, such as street, forest, or beach, etc. Accompanying the recognized place, other information based on this location can also be retrieved. Thus, such a system could be used in a variety of applications in navigation, information mining, education, games and entertainment. With the future integration with mobile technologies, such systems may feature highly in tomorrow's ubiquitous computers.

This paper briefly describes the data sampling approach required for the system as well as the feature extraction and place recognition techniques used. It also illustrates a realistic scenario of retrieving information of facilities and amenities, such as automatic teller machines, cafes, bus stops, clinics and carparks, around the vicinity of a recognized place.

MM'07, September 23–28, 2007, Augsburg, Bavaria, Germany.
ACM 978-1-59593-701-8/07/0009.

2. PLACE RECOGNITION SYSTEM
In this paper, a scene is defined as a loose collection of visual entities existing in the continuous real world. As such, a scene image is highly variable and fluid, especially in terms of the positioning of objects within the image, scale of objects and lighting conditions. In our place recognition system, we propose that the performance of the system in terms of recognition rate, can be improved by using local invariant features to handle the variances in an outdoor scene, as well as multiple queries to reduce ambiguity of user image acquisition.

2.1 Viewer-Centric Image Data Sampling
To perform place recognition, the system was trained on a set of images collected using the viewer-centric data sampling approach. In this approach, images were collected at different locations within certain distances to a specific geographical position, and placed into the same image class. This is as oppose to the object-centric whereby images are collected at different angles around an object. Hence, where a class of images using the object-centric approach contains images of the same object, the viewer-centric approach will have images of multiple objects around a geographical location. As such, a place recognition system adopts the viewer-centric approach and recognizes a place by determining the class of an input query image.

2.2 Features Extraction and Matching
There has been much progress in the development of local invariant features for image matching and object recognition. In particular, the Scale Invariant Feature Transform (SIFT) [1] is currently one of the more popular techniques used for extracting local features that are invariant to common images transformations. In our system, we adopt the SIFT as the local descriptor, due to its strength in alleviating problems that arise from variations in scale, rotation, illumination, and viewpoint. In a typical local-feature-based scene recognition system, local descriptors are extracted from each training image and stored separately in a database. When a new query image is received, local descriptors are extracted from the query images and a feature-to-feature matching is performed. The image with the closest match or its class will be returned as the output of the system

2.3 Multiple Query Images
The recognition rate of a place recognition system could be significantly improved by using multiple query images. One constraint is that in each query instance, all the query images have to be from the same class. In the real-world, this translates to the

user taking a few photographs of different surrounding objects. In our experiments, it was found that with this method, near perfect recognition could be obtained.

3. DEMONSTRATION OVERVIEW

The demonstration consists of two main modules: 1) query formulation, and 2) location and information display. For the demonstration, a dataset of 36 classes of places of images obtained from within the National University of Singapore was used. The aim of the system is to perform place recognition on an input image, display on a map its location, and provide information about amenities around it.

3.1 Query Formulation Module

For the demonstration, two different query formulation methods were used; namely: 1) query by image panorama cropping, and 2) query by existing image. Both these methods are variants of the query by example (QBE) method [2]. The first method shown in Figure 1, is a simulation of real-world image acquisition, and used only for demonstration purposes to simulate image acquisition by a camera, in a real-world scenario. In the demonstration, a user crops a portion of a panoramic image, which was generated by Autostitch™ [3], using methods described in [4]. The second method of query formulation is the query by example technique by which pre-acquired images that are stored in the database are used for querying.

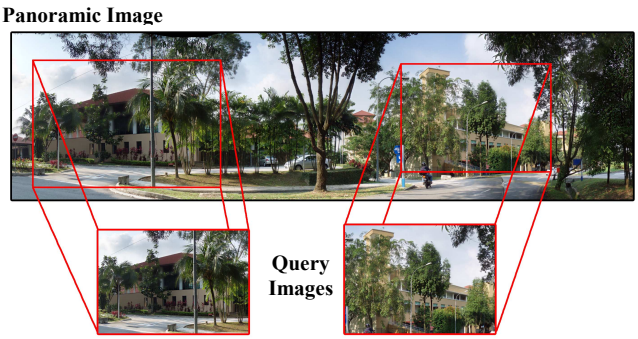

Figure 1. Query by image panorama cropping.

To improve the place recognition performance, our system caters for multiple queries to be used. Images from multiple queries may be acquired through either one or both of the above mentioned query formulation methods.

3.2 Location and Information Display Module

After the query is submitted, it is processed and through the feature matching process, the class of the image(s) is retrieved. Based on this retrieved class, information about the place can be retrieved. For the demonstration, the place is highlighted on two maps of different scales, with one illustrating its location within the campus, and one showing details of buildings and locations of amenities around the area. Further meta-information regarding the location is also retrieved, such as the names of buildings surrounding the place, and information regarding nearby amenities such as food and beverages outlets, the list of buses stopping within the area, and other facilities such as automatic teller machines and medical clinics. A screenshot of the module is shown in Figure 2.

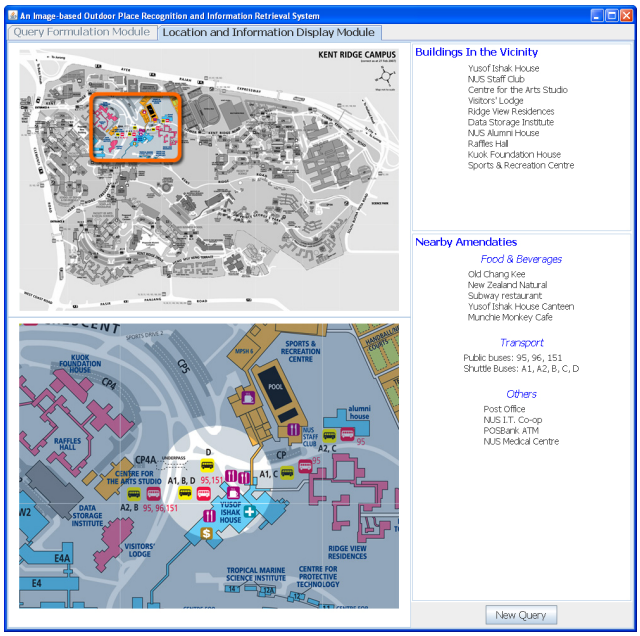

Figure 2. Location and information display module.

4. CONCLUSIONS

In this paper, we introduced a system that performs image-based outdoor place recognition by adopting the viewer-centric approach of data sampling, and using multiple queries to improve recognition performance. Our proposed demonstration system uses an innovative method to simulate real-world image acquisition by cropping from panoramic images. Finally, the place is recognized and related information is retrieved. Such systems have the potential to be used in a wide variety of applications. When integrated with mobile technology, users can instantly find out their location and other location-related information.

5. ACKNOWLEDGMENTS

The authors wish to thank M. Brown and D.G. Lowe for their demo version of Autostitch™ [3].

6. REFERENCES

[1] D. G. Lowe. “Distinctive image features from scale-invariant keypoints,” *Intl. Jnl. Computer Vision*, 60(2): 91-110, 2003.

[2] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, “Query by image and video content: the QBIC system,” *IEEE Computer*, vol. 28, no. 9, pp. 23-32, September, 1995.

[3] Autostitch™: Available online at [http://www.autostitch.net].

[4] M. Brown and D. G. Lowe. “Recognising panoramas,” *In Proc of. 9th Intl. Conf. on Computer Vision*, 1218-1225, Nice, France, October 2003.