Study Impact of Architectural Style and Partial View on Landmark Recognition

{ smileyc@stanford.edu } Ying Chen

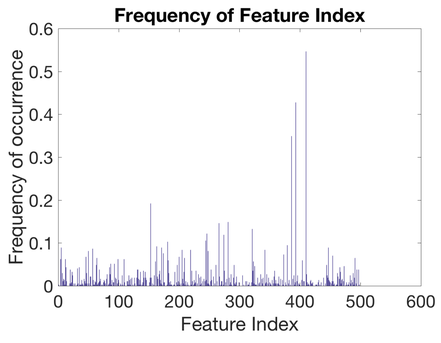
Introduction

Landmark recognition in image processing is one of the important object recognition problem. Using machine learning techniques to recognize landmarks is critical when GPS tag information is missing. Some recent studies[1][2][3] showed good accuracy of landmark detection and recognition, while most of the landmarks in the images were located in Europe and North America. In this study, the mix of Asian style and Europe/North America landmarks were used to understand how this affect recognition accuracy. The final results show that the averaged recognition accuracy is 85%.

Feature Extraction

The local features of images were detected, through finding points of interest, that represents the unique content of objects. The local feature descriptor - SURF (Speeded Up Robust Features) was used. Once all the local features of an image are computed, the image is encoded in a feature vector. It will be used as the basis for training the classifiers.

The following figure shows the frequency of each feature index of an image, which is the feature vector representation of the image.



Dataset

Google image search was used to collect the images with keywords of name of landmarks. All the images were visually checked to label the correct category. There are 18 landmarks used in this report. Nine of them are located at Asian countries, and the other nice of them are located at Europe or North America. At the training stage, the images with the global front view of landmarks were used. At the test stage, the images with partial view were also used.

Categorization

The SVM (support vector machines) model was used to categorize each test image into trained landmarks. The SVM model was described as the following equation.

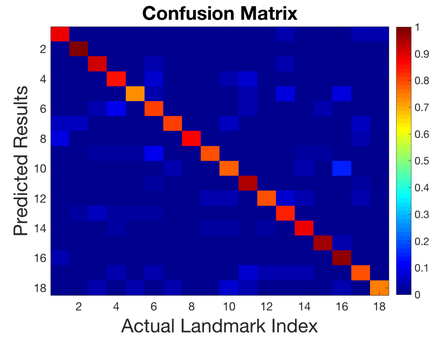
$$f(x) = sign(\Sigma_i y_i \alpha_i K(x, x_i) + b) \quad (1)$$

where x_i are the training features, y_i is the label of x_i and K is the kernel.

In this study, since multi-class problems need to be modeled, m SVMs were trained. Each SVM corresponds to each landmark category. For any test image, it was assigned to the class with largest SVM.

01	Badshahi Mosque	**02**	Dome of the Rock
03	Forbidden City	**04**	Great Wall
05	Azadi Tower	**06**	Potala Palace
07	Qutub Minar	**08**	Taj Mahal'
09	Temple of Heaven	**10**	Big Ben
11	Colosseum	**12**	Eiffel Tower
13	Golden Gate	**14**	Leaning Tower Pisa
15	Lincoln Memorial	**16**	Sagrada Familia
17	St. Basil Cathedral	**18**	Statue of Liberty

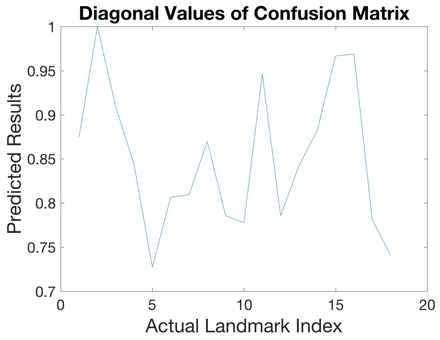
Results



The right plot only shows the diagonal values of confusion matrix of test dataset. It can be found that the recognition accuracy of most of landmarks are higher than 80%. About half of them have recognition accuracy larger than 85%, which are good results.

The averaged prediction of training dataset is 93%. The averaged prediction of test dataset is 85%. The left image shows the confusion matrix of test dataset. The highest prediction accuracy is landmark Dome of the Rock, which is 100%. The rediction accuracy is landmark Azadi Tower, which is 73%.

The landmarks with index 1 9 are located at Asian countries, which others are located at Europe or North America.



Discussions

- When the image shows the front and global view of landmark, the recognition accuracy is very good. It shows the feature detection and categorization method work very well for this task.
- When the image shows the side or partial view of landmark, the recognition error is higher. For example, the image on right has the view between front and side view.
- This issue of side or partial view may cause the feature vector of a landmark less representative of correct landmark. This is expected and may need more work to add the new feature to this landmark.

References

[1] D. Crandall, Y. Li, S. Lee and D. Huttenlocher, **Recognizing Landmarks in Large-Scale Social Image Collections**, Visual Analysis and Geolocalization of Large Scale Imagery 2016.
[2] A. Crudge, W. Thomas and K. Zhu, **Landmark Recognition Using Machine Learning**, CS229, Project 2014.
[3] Y. Zheng, M. Zhao, Y. Song, H. Adam, U. Buddemeier, A. Bisacco, F. Brucher, T. Chua and H. Neven, **Tour the World: building a web-scale landmark recognition engine**, 2009.

Future Research

As mentioned in the discussions, the partial and side view issue should be addressed to improve the overall accuracy.

In addition, how to detect landmarks with some other objects in the view is also important. For example, people in the landmark scene.