# Mining Large-Scale Social Images with Rich Metadata and Its Application

Zheng Liu[1,2], Hua Yan[1,2], Huijian Han[2]
1 School of Computer Science and Technology, Shandong Economic University, Ji'nan 250014, China
2 Shandong Provincial Key Laboratory of Digital Media Technology, Ji'nan 250014, China

Email: Liuzheng48@gmail.com

***Abstract*—In this paper, we study on how to automatically mine landmarks from large-scale social images with rich metadata. Firstly, location name is submitted to social image community, and then related social images with rich metadata are obtained. Afterwards, these social images are clustered according to different kinds of metadata of images, and candidate landmarks are mined from the images clustering results. Next, noisy landmarks are pruned from candidate landmarks by computing geographical entropy and time entropy. Experiments conducted on Flickr photos demonstrate the effectiveness of the proposed approach and our approach can also provide useful information for tourists to make tourist plans.**

***Index Terms*—Landmarks, Social Images, Co-clustering, Metadata Similarity.**

## I. INTRODUCTION

Landmark can be considered as a place where most tourists prefer to sight-see, therefore, finding landmarks of a given city could provide helpful information for tourists. Landmarks have the following two characteristics which are 1) The more popular a landmark is, the larger the number of tourists and hence the number of photos uploaded to Web. 2) Most geo-tags of a landmark should be aggregated at a small number of places.

With the rapid development of Web 2.0 techniques, there has been an increasingly large amount of online communities available on the Web. These social communities successfully facilitate the information generating, sharing, and distributing among different users. In summary, effectively finding landmarks from social images could enhance the performance of tourist guiding systems.

In the past few years, there are some pioneering works concerning on landmarks which are listed as follows.

Gao et al.[1] present a travel guidance system named W2Go, which can automatically recognize and rank the landmarks for travelers. In this paper, a novel Automatic Landmark Ranking (ALR) method is proposed by utilizing the tag and geo-tag information of photos in Flickr and travelling information from Yahoo Travel Guide. R. Abbasi et al.[2] proposed a method to identify landmark photos using tags and social Flickr groups without depending on GPS coordinates for these photos. The information they used only are Flickr tags and user groups information. They apply a SVM classifier for which the training data is extracted from Flickr groups to find relevant landmark-related tags. L. S. Kennedy et al.[3] used both context- and content-based methods to generate representative sets of images for location-driven features and landmarks. They use location and other metadata, such as tags associated with images, and the images' visual features to solve this problem. They present an approach to extract tags which could represent landmarks.

In a recent research work[4], Zheng et al. worked on the landmark recognition. They built a web-scale landmark recognition engine exploiting 20 million GPS-tagged photos of landmarks together with online tour guide systems. Adrian Popescu et al.[5] developed a system named MonuAnno which automatic annotate geo-referenced landmarks images. The proposed system exploits both image localization information and image visual content analysis. In paper [6], the authors presented a prototype system which can automatically generate landmarks using pedestrian navigation directions from geo-tagged photos. Both navigation images selecting and images with directional instructions enhancing are executed automatically. W. Chen et al.[7] present a novel data-driven approach which depends on online photos sharing websites, such as Flickr, for automatically generating tourist maps. The proposed algorithm uses the geographical areas as input and then finds geo-tagged photos from online photo collections. The algorithm generates a set of points of interest for the area by clustering the photos based on their locations and identifying the popular tags for each cluster.

## II. FRAMEWORK OF THE PROPOSED METHOD

The framework of our approach is made up of six steps which are shown in Fig.1. Given a location name, we submit it to social image community, and then obtain related images. We cluster these social images by different kinds of metadata. Next, candidate landmarks are extracted from descriptive terms of social image clusters. Afterwards, we refine candidate landmarks by geographical entropy and time entropy to generate final landmarks. Example of a social image with rich metadata is shown in Table.1.

© 2012 ACADEMY PUBLISHER
doi:10.4304/jsw.7.4.749-756

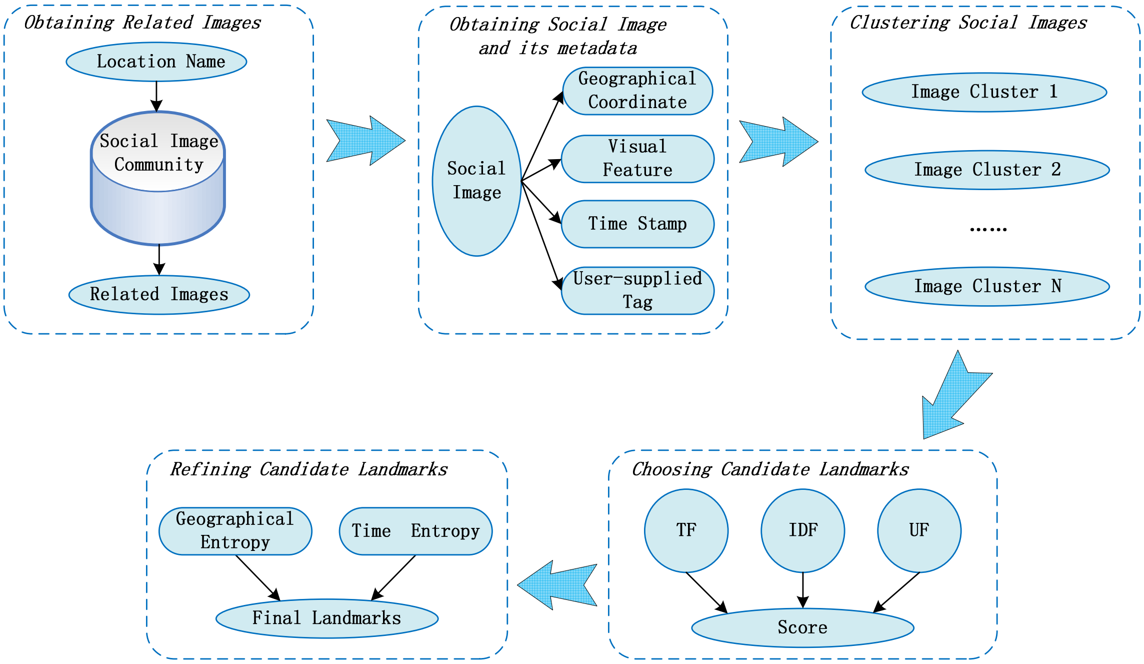



Figure 1. Framework of the proposed approach.

TABLE I. EXAMPLE OF A FLICKR PHOTO WITH METADATA.

| | |
|---|---|
| **URL** | http://www.flickr.com/photos/bruce_mcadam/3841590751/meta/in/set-72157619915025563 |
| **Author name** | Bruce McAdam |
| **Photo** |  |
| **User-defined tags** | China, Beijing, Forbidden City |
| **GPS coordinate** | Latitude: 39 deg 54' 14.39" N<br>Longitude: 116 deg 25' 20.49" E |
| **Time stamp** | Date: 2009.06.17, Time: 03:00:49 |
| **Geopolitical locations** | Beijing, China |

## III. SOCIAL IMAGES CO-CLUSTERING BY DIFFERENT KINDS OF METADATA.

The rich metadata is one of the important features for social images. If the rich metadata of social images are used fully, many social image based applications could be developed well. Social image's metadata we considered includes time stamp, GPS coordinate, user-supplied tags set and visual features, and the metadata similarity measuring methods are listed as follows.

### *A. Time Stamp Distance*

Introducing time stamp into photo clustering process can enhance the clustering performance, and the reason lies in that a small time interval may suggest that the photos were taken in the same place. When a photographer takes pictures, the surrounding scene may be fairly similar if the time interval between two photos is fairly short. For example, if one took a photo in a "beach" scene, it is impossible that another photo taken within five minutes would be in a "city" scene.

Hence, we employ the time stamp as metadata in clustering process. The less the time stamp change, the more likely the photos could be arranged into one cluster. The time stamp distance is computed as follows.

$$D_T(t_i,t_j)=\begin{cases} dis(t_i,t_j)/\xi & dis(t_i,t_j)<\xi \\ 1 & otherwise \end{cases} \quad (1)$$

where $\xi$ is a threshold to judge whether two time stamps are close to each other.

© 2012 ACADEMY PUBLISHER

### B. GPS Distance

The GPS coordinate could be represented as a two-dimensional vector which includes latitude and longitude. By GPS features, GPS distance is computed by the distance between the locations where the photos were taken. Intuitively, if the distance between the locations where two photos were taken is too large, it is impossible to group the two photos in one cluster. Following the method in [8], we define GPS distance to estimate the location relationship between two photos as follows.

$$D_G(g_i, g_j) = \log\left(R\varphi\right) \quad (2)$$

$$\varphi[rad] = 2\arcsin\left(\sqrt{\sin^2\frac{\Delta\delta}{2} + \cos\delta_{g_i}\cos\delta_{g_j}\sin^2\frac{\Delta\lambda}{2}}\right)$$

(3)

where $\delta_{g_i}$ and $\delta_{g_j}$ are the latitudes of two photos of $g_i$ and $g_j$ respectively, $\Delta\delta$ and $\Delta\lambda$ are the differences of latitude and longitude of photo $g_i$ and $g_j$. $R$ is the radius of the earth, which is equal to 6370 km.

### C. Similarity between User-supplied Tags Set

An important service Flickr provided is that users could manually annotate their photos using so called tags, which describe the contents of the photo or provide additional contextual and semantic information. Then, for a Flickr photo, a tag set can be extracted from the corresponding Web page. Therefore, the semantic relevance of tag sets could represent the relationship between photos.

We define a method named NFD to compute semantic similarity between tags which is analogous to NGD[9], which is a distance function between two words obtained by searching a pair of words using the Google search engine. NFD between two tags can be estimated based on Flickr as follows.

$$NFD(w_i, w_j) = \frac{\max\left\{\log f(w_i), \log f(w_j)\right\} - \log f(w_i, w_j)}{\log G - \min\left\{\log f(w_i), \log f(w_j)\right\}}$$

(4)

where $w_i$ and $w_j$ represent the two tags in consideration. $f(w_i)$ and $f(w_j)$ are the numbers of images containing tag $w_i$ and tag $w_j$ respectively, which can be obtained by performing search by tag on Flickr website using the tags as keywords. $f(w_i, w_j)$ is the number of the images returned by Flickr when typing $w_i$ and $w_j$ as the search term respectively. Furthermore, $G$ is the total number of images in Flickr. The concurrence similarity between tag $w_i$ and tag $w_j$ is then defined as.

$$\gamma(w_i, w_j) = \exp[-NFD(w_i, w_j)] \quad (5)$$

The related tag set of photo $p$ and $q$ could be represented as $S^p = \{w_1^p, w_2^p, \ldots, w_m^p\}$ and $S^q = \{w_1^q, w_2^q, \ldots, w_n^q\}$. The similairty between tag set $S^p$ and $S^q$ is computed as follows.

$$D_W(S^p, S^q) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\gamma(w_i^p, w_j^q)}{m \cdot n} \quad (6)$$

### D. Image Visual Similarity

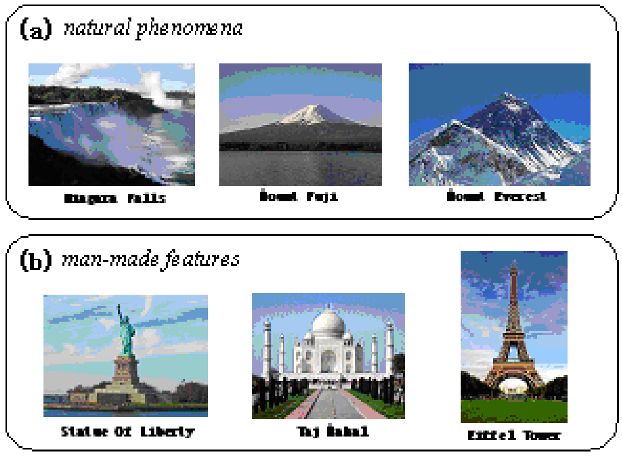



Figure 2. Illustration of two types Landmarks.

As is shown in the Web page[1] of Wikipedia, we know that landmarks can be split into two categories: *natural phenomena* (physical features such as waterfalls, rivers, mountains and rock formations) and *man-made features* (like buildings, bridges, statues, public squares and so forth). This list includes natural phenomena by country, followed by man-made features by continent and by country. In the image content analysing field, both global features and local features can be applied. However, these two features specialize in different cases. In a word, global features have the ability to generalize an entire object with a single vector. Consequently, their use in standard classification techniques is straightforward. Local features, on the other hand, are computed at multiple points in the image and are consequently more robust to occlusion and clutter[10].

Considering the different application scenarios of global and local features, we use both of them to deal with landmark photos. When the landmark belonging to a *man-made features* category, it is possible to find salient objects in related photos. Hence, in this case, local features could perform better than global features. On the other hand, if the landmark to be detected is belonged to *natural phenomena* class, global features are more important. Therefore, we introduce both global and local features in our system, and dynamic tune the weight of them to enhance the image content analysis capability. As is shown in Fig.2, there may be buildings or salient objects in landmark photos which belong to the *man-made features* category.

We totally extracted 168-dimension color and texture features (shown in Table.2) as the low-level visual representation of the images. In addition, we employ cosine similarity to estimate the visual similarity between a pair of images based on global features(Shown in Eq.7).

[1]http://en.wikipedia.org/wiki/List_of_landmarks

© 2012 ACADEMY PUBLISHER

TABLE II. THE LOW-LEVEL FEATURES EXTRACTED FROM IMAGES.

| Feature category | Feature Name | Dimensions |
|---|---|---|
| Color | Color Correlogram | 44 |
| | Color Texture Moment | 14 |
| | Color Moment | 6 |
| Texture | Wavelet Features [11] | 104 |

$$Sim_G(I_i, I_j) = \frac{v_i \cdot v_j}{\|v_i\|\|v_j\|} \quad (7)$$

where $v_i$ and $v_j$ are the global feature vectors of $I_i$ and $I_j$ respectively.

Inspired by the recent progress in object recognition, we use the visual word model and SIFT features[12] to measure image similarity. We use SIFT to describe the regions around the keypoints. To construct codebook, we use corel5k dataset[13] as training data. All SIFT descriptors in the image of corel5k dataset are grouped into clusters which are named visual words by vector quantization with the Linde-Buzo-Gray (LBG) algorithm[14], which is a vector quantization algorithm to derive a good codebook. After vector quantization, all images are represented as a *D*-dimensional vector, and the value of *D* is equal to the number of visual words. Then, each image can be represented as a histogram, and each histogram bin corresponds to a visual word. Using the histograms as feature vectors representing the images, image visual similarity can be computed by the distance between feature vectors. In our experiments, 2000 visual words are used for all photos.

Supposing that visual words vector $h_i$ and $h_j$ are *D*-dimensional (*D* equals to the vocabulary size of visual words) vectors of visual word frequencies, which come from image $I_i$ and image $I_j$ respectively. Then the image similarity based local feature is computed as follows.

$$Sim_L(I_i, I_j) = \frac{h_i^t h_j}{\|h_i\|\|h_j\|} \quad (8)$$

Then, combining global features and local features, the overall image similarity can be obtained by linearly combining both global and local features as follows.

$$Sim(I_i, I_j) = \alpha \cdot Sim_G(I_i, I_j) + (1-\alpha) \cdot Sim_L(I_i, I_j), \ \ 0 < \alpha < 1 \quad (9)$$

As is shown in Eq.9, the parameter $\alpha$ is used to adjust the influence of global and local feature on image similarity measuring. Intuitively, in the case of "*natural phenomena*" landmark photos, $\alpha$ should be larger than 0.5, otherwise, $\alpha$ is set smaller than 0.5 when considering "*man-made features*" landmark photos.

### *E. Social Images Co-clustering Algorithm*

As there is a large amount of online visual information in photo shared community, the large-scale Web images should be pre-clustered based the on the rich metadata to reduce the computational cost. How to effectively integrate different kinds of metadata is a key issue in social image clustering problem. Inspired by the idea of image and low-level feature co-clustering[15], we use a k-partite spectral graph model to describe the relations among visual features, user-supplied tags, GPS information and other kinds of metadata. A k-partite graph is defined as a graph whose graph vertices can be partitioned into k disjoint sets so that no two vertices within the same set are adjacent. To tackle the k-partite spectral graph partitioning problem, we modify the approach proposed in [16][17] and then present a k-partite graph partitioning model to describe the relations between different kinds of metadata of the social images.

The k-partite graph could be partitioned by the method proposed in [15] which is named consistent bipartite graph co-partitioning (CBGC). CBGC mainly concerns about the consistent fusion of two co-clustering sub-problems. Spectral graph partitioning is an effective heuristic that was introduced in the early 1970s, and popularized in 1990. Spectral partitioning generally gives better global solutions than the KL or FM methods[18].

To describe our graph-based co-clustering model, we give an example to illustrate how to partition a tripartite graph as shown in Fig. 3.

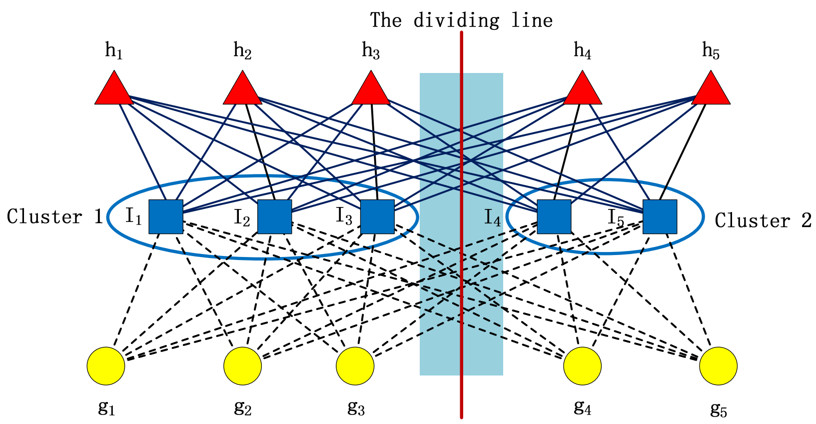



Figure 3. Illustration of partitioning the tripartite graph to cluster images.

© 2012 ACADEMY PUBLISHER

where $I_i$ denotes as the *i*-th image, moreover, $g_j$ and $h_k$ denote two different kinds of metadata points respectively. The edge weight of a k-partite spectral graph is represented by metadata similarity (or distance) which has been explained in section 3.

The key problem of our spectral graph partitioning problem is to efficiently optimize a specific objective function. Specially, CBGC only focus on a partition of two clusters, where all types of objects will be simultaneously clustered into two groups respectively. To extend the number of clusters, we could simply perform the above bipartition algorithm in a recursive mode.

## IV. Finding Landmarks From Social Images

For a given location, such as a city, we firstly use the algorithm proposed in section 3 to co-cluster the photos which are returned by submitting the name of location to a social image community(such as Flickr), and then a large number of images related to the specific location can be returned. We then choose descriptive terms for landmarks from the textual information of social image clusters and then remove noisy landmarks from candidate landmarks.

### *A. Finding Candidate Landmarks' Description from Social Image Clusters*

Our approach is based on the TF/IDF technique which is widely used in information retrieval and text mining. TF/IDF weight is a statistical measure used to evaluate how important a word is to a document in a collection or corpus. The importance increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word in the corpus. We assume that the user-supplied tags of Flickr photos captured in a photo cluster and do not frequently occur in other clusters are more representative to a landmark.

In [19], Rattenbury et al. proposed a TF/IDF-like approach to determine whether each tag has a coherent description of places. Rattenbury's approach is also based on the TF/IDF technique. In this paper, we follow the method proposed in paper [19] to filter out excessively common terms and choose representative terms.

To judge whether tag $t$ could represent a landmark or not, we introduce three factors to score tag $t$ based on the traditional TF/IDF calculation. We define term frequency for a given tag $t$ within a cluster $C$ is the count of the number of times $t$ was used in $C$, the computing method is as follows.

$$tf(C,t) \triangleq num(\mathrm{P}(C,t)) \tag{10}$$

where $\mathrm{P}(C,t)$ represents the photos belonged to cluster $C$ and tagged by $t$ as well. The inverse document frequency for tag $t$ computes the overall ratio of the tag $t$ among all photos.

$$idf(C,t) \triangleq \frac{num(\mathrm{P})}{num(\mathrm{P}(t))} \tag{11}$$

We believe that a tag is more representative if there are a large number of different users use it in a cluster. Specifically, we compute the percentage of the photographers in cluster $C$ who use tag $t$ (shown in Eq.12).

$$uf(C,t) \triangleq \frac{num(U(C,t))}{num(C)} \tag{12}$$

where $U(C,t)$ represents the photographers in cluster $C$ who used tag $t$. Combining the above three factors, the final score for tag $t$ in cluster $C$ is computed as follows.

$$S(C,t) = tf(C,t) \cdot idf(t) \cdot uf(C,t) \tag{13}$$

We believe the higher the score that a term get, the more likely it could be a name of landmark. Afterwards, we choose the terms with higher score as the candidate landmarks.

### *B. Candidate Landmarks Refinement*

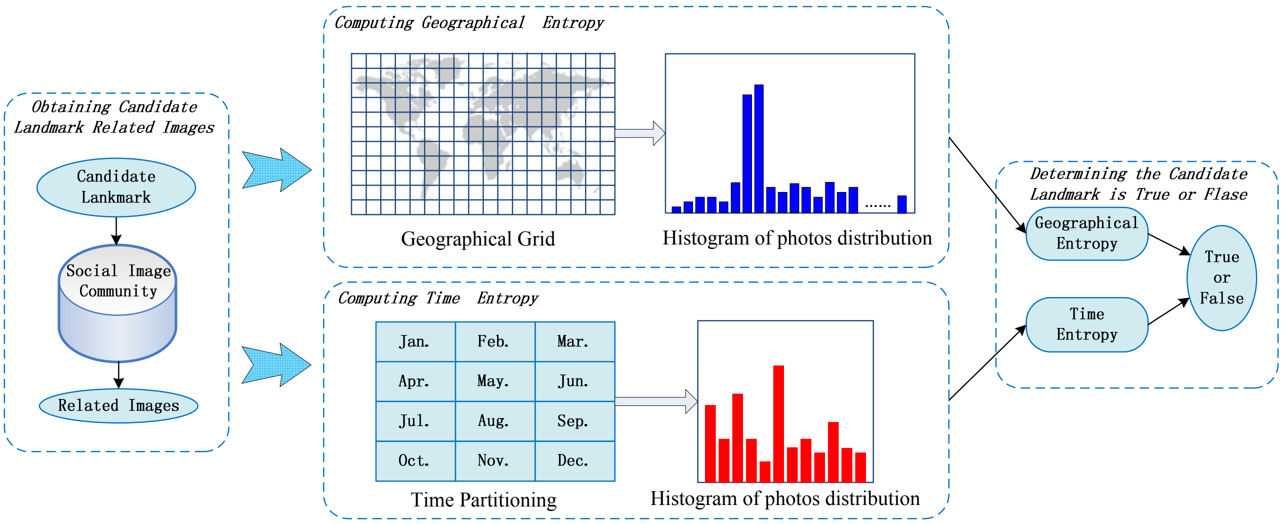



Figure 4. Illustration of candidate landmarks refinement.

© 2012 ACADEMY PUBLISHER

However, the candidate landmarks may be noisy. From some experiment results, we find that some hot events or activities may be falsely recognized as landmarks from geo-clusters. For instance, after Beijing 2008 Olympic Games, a great many of photos related to it were uploaded to social image community. Particularly, Beijing Olympic Swimming game was hosted in National Aquatics Center, and the photos taken at this swimming game converge at a limited area. Hence, it is possible to falsely deem the event “Beijing Olympic Swimming” as a landmark. Based on the above analysis, we could detect the landmark not only using geographical information, but also adopting the time information when the candidate landmark photos were taken. The geographical entropy of a candidate landmark $L$ with $n$ regions $\{r_1, r_2, \cdots r_n\}$ on the earth is defined as follows.

$$H_{GE}(L) = -\sum_{i=1}^{n} p(r_i) \cdot \log_2 p(r_i) \tag{14}$$

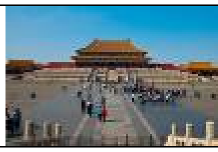
where $p(r_i)$ is the probability that the photos related to the candidate landmark are taken in the region $r_i$, and $p(r_i)$ can be simply obtained by Eq.15.

$$p(r_i) = \frac{Num(r_i)}{\sum_{j=1}^{n} Num(r_j)} \tag{15}$$

where $Num(r_i)$ is the number of photos taken in region $r_i$.

Afterwards, threshold $\xi$ could be set to prune noisy landmarks, that is, the landmark $L$ is reserved as true one only if $H_{GE}(L) < \xi$.

TABLE III. PERFORMANCE OF LANDMARK RECOGNITION BY TWO DIFFERENT METHODS.

| City | Approach | Rank 1# | Rank 2# | Rank 3# | Rank 4# |
|---|---|---|---|---|---|
| Beijing | YTG | Great Wall | Lama Temple (Yonghegong) | Forbidden City | Temple of Heaven |
| | Our Approach | Great Wall | Forbidden City | Beijing National Stadium | National Aquatics Centre |
| | Typical Image |  |  |  |  |
| Shanghai | YTG | Oriental Pearl TV Tower | Yu Yuan Gardens | Shanghai Urban Planning Exhibition Hall | Nan Jing Road |
| | Our Approach | Oriental Pearl TV Tower | China Pavilion | Jinmao Tower | People's Square |
| | Typical Image |  |  |  |  |
| Hongkong | YTG | Victoria Peak | Ocean Park | Victoria Harbour | Nathan Road |
| | Our Approach | Victoria Harbour | International Finance Center | Bank of China Tower | Hongkong Disneyland |
| | Typical Image |  |  |  |  |

On the other hand, some hot events may also be centered in a specific small area. Therefore, we define time entropy to prune noisy landmarks. Similar to geographical information entropy, we partition each year to every months and compute time entropy for each candidate landmark(shown in Eq.16).

$$H_{TE}(L) = -\sum_{j=1}^{12} p(s_j) \cdot \log_2 p(s_j) \tag{16}$$

© 2012 ACADEMY PUBLISHER

where $p(s_j)$ is the probability that the photos related to the candidate landmark are taken in the month $j$. We set a threshold $\delta$ to remove noisy landmarks if $H_{TE}(L) > \delta$. In our experiments, parameter $\xi$ and $\delta$ are set to 3.5 and 1.8 respectively.

## V. Experimental Results and Analysis.

In this section, we evaluate the performance of our approach for mining landmarks from Flickr photos. To evaluate the performance of our approach, we compare it with a baseline method which is the Yahoo Travel Guide(YTG)[20]. Yahoo Travel Guide provides an area based guide service. In each country, several main cities are listed. For example, there are a total of 10 cities listed for China, and they are Shanghai, Hongkong, Beijing, Guangzhou, etc. For each city, Yahoo Travel Guide provides information on key attractions where a series of ranked landmarks are shown to users, associated with comments by users. For each city, we crawl geo-tagged images from Flickr with tags and metadata using the Flickr API[21]. We design two experiments schemes to test the performance of the proposed approach. In experiment 1, we compare the landmark detecting results for three cities using different methods(shown in Table.3).

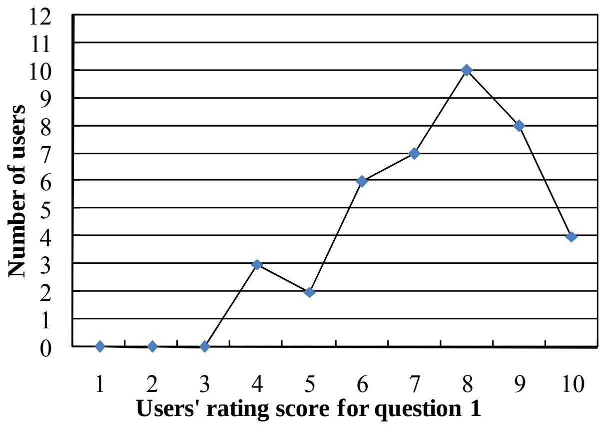



Figure 5. Evaluation on question 1.

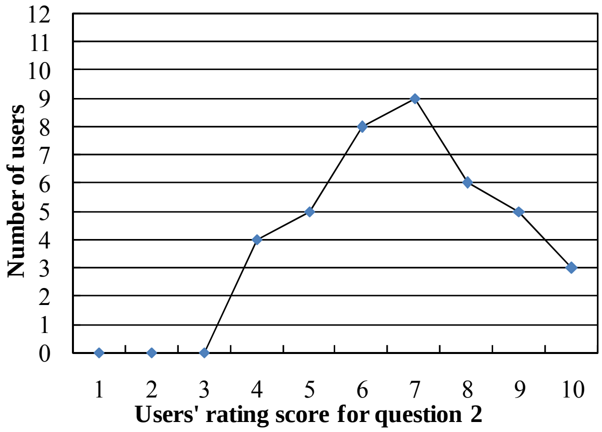



Figure 6. Evaluation on question 2.

To value the performance of our approach, we arrange 40 volunteers to value the proposed approach by providing the landmarks of a given city detected by our approach. We ask each participant to answer two questions by providing a rating score of between 1(very bad) and 10(excellent): 1) *Question 1# is if our approach could find the most important landmarks of a city?* 2) *Question 2# is if the landmarks detected by our approach are ranked accurately?* As is shown in Fig.5 and Fig.6, the average rating score of the above two questions are 7.48 and 6.88 respectively. Hence, we can see that our approach could provide helpful information for users.

## VI. Conclusions and Future Works

This paper proposes an automatic landmark mining method by social images clustering. Our approach only need the users to provide name of a city, and then photos and related metadata are crawled from Flickr website. According to the rich metadata of photos, we could cluster these Flickr photos. Finally, we can detect landmarks of the given city from photo clusters.

In the future, we would extend our work in the following aspects. Firstly, we will try other social image community to examine the performance of our approach. Secondly, we will attempt to modify our method into parallel model to make it more suitable for massive social image dataset.

## Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant No.60970048) and Project of Shandong Province Higher Educational Science and Technology Program (Grant No.J10LG69). We also appreciate the anonymous reviewers for thoroughly reading the paper and providing thoughtful comments.

## References

[1] Yue Gao; Jinhui Tang, Richang Hong, Qionghai Dai, Tat-Seng Chua, Ramesh Jain. W2Go: A Travel Guidance System by Automatic Landmark Ranking. In MM '10: Proceedings of the seventeen ACM international conference on Multimedia, 2010.

[2] R. Abbasi, S. Chernov, W. Nejd, R. Paiu, and S. Staab. Exploiting flickr tags and groups for finding landmark photos. ECIR, 2009.

[3] Lyndon Kennedy and M. Naaman. Generating diverse and representative image search results for landmarks. WWW 2008.

[4] Y. Zheng, M. Zhao, Y. Song, H. Adam, U. Buddemeier, A. Bissacco, F. Brucher, T. Chua, and H. Neven. Tour the world: building a web-scale landmark recogntion engine. CVPR, 2009.

[5] Adrian Popescu, Pierre-Alain Moëllic. MonuAnno: Automatic Annotation of Georeferenced Landmarks Images, CIVR 2009, July 8-10, Santorini, Greece.

[6] H. Hile, R. Vedantham, A. Liu, N. Gelfand, G. Cuellar, R. Grzeszczuk, and G. Borriello, Landmark-Based Pedestrian Navigation from Collections of Geotagged Photos, in MUM 2008: Proceedings of ACM International Conferenceerence on Mobile and Ubiquitous Multimedia. ACM Press, 2008.

[7] W. Chen, A. Battestini, N. Gelfand, and V. Setlur. Visual summaries of popular landmarks from community photo collections. In ACM international conference on Multimedia, pp. 789-792, 2009.

[8] Yuki Arase, Xing Xie, Takahiro Hara, Shojiro Nishio. Mining People's Trips from Large Scale Geo-tagged Photos. ACM MM2010, pp.133-142.

[9] Rudi L. Cilibrasi, Paul M.B. Vitányi. The Google Similarity Distance, IEEE Trans. on Knowledge and Data Engineering. 19(3), pp.370-383, 2007.

[10] D. A. Lisin, M. A. Mattar, M B. BlMark C. Benfield and E. G. Learned-Miller. Combining Local and Global Image Features for Object Class Recognition. In CVPR, 2005.

[11] Chang, T., and Kuo, C.C.J. Texture analysis and classification with tree-structured wavelet transform, IEEE Trans. on Image Processing, vol. 2, no. 4, pp.429-441, 1993.

[12] D. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60(2):91-110, 2004.

[13] Duygulu, P., Barnard, K., de Freitas, J., Forsyth, D.A. Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary. Proc. 7th Eur. Conf. on Computer Vision:97-112, 2002.

[14] Linde, Y., Buzo, A., Gray, R. An Algorithm for Vector Quantizer Design, IEEE Transactions on Communications, 28:84-94, 1980.

[15] Bin Gao, Tie-Yan Liu, Tao Qin, Xin Zheng, Qian-Sheng Cheng, Wei-Ying Ma. Web image clustering by consistent utilization of visual features and surrounding texts. MM2005, pp.112-121., 2005.

[16] Bin Gao, Tie-Yan Liu, Qian-Sheng Cheng, Guang Feng, Tao Qin, and Wei-Ying Ma, Hierarchical Taxonomy Preparation for Text Categorization Using Consistent Bipartite Spectral Graph Copartitioning, IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 9, pp. 1263-1273, September 2005.

[17] Bin Gao, Tie-Yan Liu, Xin Zheng, Qian-Sheng Cheng, and Wei-Ying Ma, Consistent Bipartite Graph Co-Partitioning for Star-Structured High-Order Heterogeneous Data Co-Clustering, in Proceedings of the 11th International Conference on Knowledge Discovery and Data Mining, pp. 41-50, 2005.

[18] Dhillon, I.S. Co-clustering documents and words using bipartite spectral graph partitioning. KDD'01, 2001.

[19] Rattenbury, T. and Naaman, M. Methods for Extracting Place Semantics from Flickr Tags, In ACM Trans. on the Web, Vol. 3, No. 1, pp. 1-30, 2008.

[20] http://travel.yahoo.com/.

[21] http://www.flickr.com/services/api.



**Zheng Liu**, born in 1980, earned a B.S. and M.S. degree in Computer Science & Technology from Shandong University, in 2002 and 2005 respectively. After graduate school, he joined School of Computer Science & Technology, Shandong Economic University in 2005. Now he is pursuing his Ph.D. degree in Shandong University, and works with the Information Retrieval Group. His current research interests include machine learning, pattern recognition and multimedia data mining.

**Hua Yan**, born in 1973, received the Ph.D. degree in communication and information system from Shandong University, Jinan, China, in 2007. Now she is an associate professor in Shandong Economic University. Her current research interests include various aspects of image/video processing and its application.

**Huijian Han**, Male, was born in HeZe city, China, in December 19, 1971. In 2000, Han is a master majoring in Computer Software Theory and Techniques of Shandong University, China, and now as Computer Applied Technique Ph. D. candidate in Shandong University. Han's major field of study is texture mapping of CG. He is a PROFESSOR, MASTER TUTOR in School of Computer Science & Technology of Shandong Economic University and Shandong Prov. Key Lab of Digital Media Technology, in Jinan city, China. He participated in the work in 1992 and has long been engaged in computer technology in teaching and research work. He has published more than 30 papers in academic journals at home and abroad and participated in the preparation of two books. For example: Computer Graphics (Beijing, China: Science Press,2005); Concise Guide to Computer Graphics (Beijing, China, Higher Education Press,2007); Determining Knots by Minimizing Energy(Beijing , China, Journal of Computer Science and Technology, 2006). His current research interests include CG&CAGD, computer simulation and digital media techniques.

Prof. Han is a member of Chinese Association for System Simulation, Computer Society of Shandong Province, and Academic Committee of Shandong Prov. Key Lab of Digital Media Technology. Prof. Han received third prize of Shandong Province Natural Science Award in 2005, received first prize of Science and Technology Progress Award of People's Republic of China Ministry of Education in 2007 and received Outstanding Contribution Award from the Shandong Provincial Science and Technology Association in 2008.


© 2012 ACADEMY PUBLISHER