

Discovering Obscure Sightseeing Spots by Analysis of Geo-tagged Social Images

Chenyi Zhuang, Qiang Ma, Xuefeng Liang, Masatoshi Yoshikawa
 Graduate School of Informatics, Kyoto University,

Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, JAPAN

Email: zhuang@db.soc.i.kyoto-u.ac.jp, {qiang, xliang, yoshikawa}@i.kyoto-u.ac.jp

Abstract—In contrast to conventional studies of discovering hot spots, by analyzing geo-tagged images on Flickr, we introduce novel methods to discover obscure sightseeing spots that are less well-known while still worth visiting. To this end, we face two new challenges that the classical authority analysis based methods do not encounter: how to discover and rank spots on the basis of 1) popularity (obscurity level) and 2) scenery quality. For the first challenge, we estimate the obscurity level of a spot in accordance with the visiting asymmetry between photographers who are familiar with a target city and those who are not. For the second challenge, the behavior of both viewers who browsed the images and photographers are analyzed per each spot. We also develop an application system to help users to explore sightseeing spots with different geographical granularities. Experimental evaluations and analysis on a real dataset well demonstrate the effectiveness of the proposed methods.

I. INTRODUCTION

Benefiting from Social Networking Service (SNS) and advances in mobile devices, people can upload and share their experiences on the Internet. Geo-tagged images are one of the typical content generated by users. The vital information it contains provides researchers with excellent opportunities for discovering landmarks and moving patterns. For instance, GPS traces [1], images [2], check-ins [3], and tweets [4] are treated as different kinds of user votes to help gather tourism knowledge. Authority based analysis, like “rank-by-count” and “rank-by-frequency” in a vote manner, is the basis for most of these trip recommendation research.

More importantly, based on a survey by Zheng et al. [5], the growing geo-referenced and community-contributed media resources have generated huge amounts of detailed location and event tags, covering not only famous landmarks but also obscure locations. Because obscure locations always have not enough visits or votes on the Internet, the conventional authority based analysis, which is used to recommend popular locations, is not suitable.

By dividing sightseeing spots into four quadrants on the basis of their “popularity” and “sightseeing quality”, we are trying to discover the obscure ones [6]. Located in the quadrant with high sightseeing quality but low popularity, an obscure sightseeing location is a considerable choice for in-depth travel to not only enjoy the beautiful scenery but also experience local culture, especially for the repeat tourists who have already visited the most famous places there.

In this paper, to a given scenery object, we propose a system to discover the obscure sightseeing locations with two novel

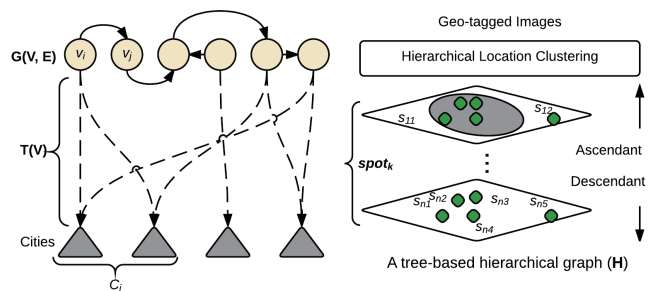


Fig. 1. Data construction

metrics of locations, “obscurity level” and “scenery quality”, by analysis of geo-tagged images retrieved from Flickr [7]. In section III, we first devise novel methods to classify Flickr users by considering their familiarity with a given city. In contrast to conventional methods, we pay attention to users’ properties and classify them into different groups to detect potential valuable spots. Then, we evaluate scenery quality of a spot by analyzing user behavior related to the images shot around it. The behavior of both viewers who browsed the social images (i.e., social appreciation) and photographers (i.e., photographers’ attention) are considered in section IV. Implementations of the system are validated on the basis of field visits in Section V.

II. OVERVIEW OF OUR SYSTEM

This section introduces our system, which is comprised of three components: data construction, obscure sightseeing spots mining, and recommendation.

A. Data Construction

Figure 1 illustrates the relationships between the main three data structures in our system.

Definition 1: A social network G is a directed graph $G = G(V, E)$, where V is the photographer set of v_i and E is the friendship set of $e(v_i, v_j)$ from v_i to v_j .

Definition 2: A familiarity matrix T_i is a mapping between a photographer v_i and all the cities C_i that s/he has been to. For example, $T_i[y = 2014, c = kyoto] = 10, c \in C_i$ means that among v_i ’s images taken in Kyoto in 2014, there are 10 different dates detected from the images’ taken dates.

Definition 3: A tree-based hierarchical graph H is a collection of location clusters, denoted as $spot_k$, with a geographical hierarchical structure. A $spot_k$ is a candidate sightseeing place, and all the images taken there are denoted as an image set I_k .

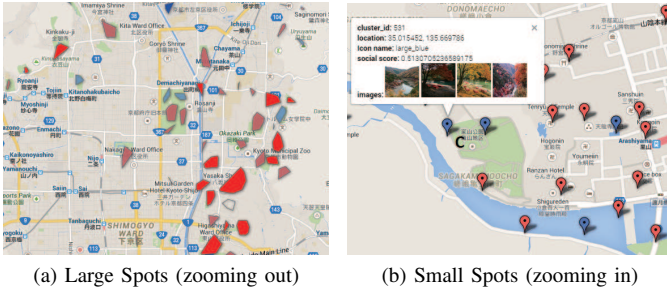


Fig. 2. Exploring interfaces

B. Obscure Sightseeing Spots Mining

Obscurity Level Estimation: Given $G(V, E)$ and T_i , we first estimate each photographer v_i 's familiarity with a particular city $c \in C_i$, based on which all the photographers are classified into a *familiar group* and an *unfamiliar group*. Then, on the basis of visiting frequency asymmetry between the photographers from these two groups, we estimate the obscurity level of each $spot_k$.

Scenery Quality Evaluation: Two kinds of user behavior are identified to evaluate the scenery quality. Regarding the behavior of users who viewed target images, we evaluate a $spot_k$'s social appreciation to present how much on-line community members like some scenery around there. On the other hand, by analyzing the behavior of photographers who took target images, we try to evaluate the attention they have paid to some scenery at $spot_k$.

C. Recommendation

Given a query, say “Kyoto, Maples”, the recommendation component first labels all the related geo-tagged images (see Section IV-A). Then, as shown in Figure 2, a list of ranked sightseeing spots will be displayed on a Web map. By changing the zoom levels, a user can explore the spots in different geographical granularities. From deep blue (obscure) to deep red (popular), different colors indicate the obscurity levels of maple spots. By clicking any tags on the map, a user can view the images taken there, as well as the ranking information.

III. OBSCURITY LEVEL ESTIMATION

A. Photographer Classification

In this subtask, we will first introduce two complementary familiarity scores, F_s and F_v , to classify photographers into the *familiar group* and *unfamiliar group*. Then, the relationship between these two scores will be discussed.

1) *Social Network Based Familiarity Score:* If a photographer is a resident, we assume s/he is familiar with the target city. However, few users specify their residency information on Flickr. Due to this, we propose a social network based method to calculate the familiarity of a photographer with a given city c . We assume that a photographer with many friends who are familiar with city c will also be familiar with the city c .

$$F_s(c) = \alpha_B \cdot G \cdot F_s(c) + (1 - \alpha_B)d_c. \quad (1)$$

By iterating the biased PageRank on the graph G , the familiarity value of a parent is split among its children, and

the familiarity value of a child node is the sum of familiarity values propagated over its links. This is exactly consistent with our assumption. For the normalized bias distribution d_c , compared with the TruskRank [8], our system can automatically select good seeds due to whether a v_i is a resident or not as detected from his/her profile. The entries of the vector d_c that correspond to good seeds sum up to 1. Given a city c , $F_s(c)$ is initialized to d_c before the iteration of the above computation. α_B is the decay factor.

Xu et al. [9] have also proposed a similar assumption and proved that geographical nearby users are more probable to establish friendship relations using their dataset. However, the sensitivity (also known as recall) of this method would become low because of the exception that a potential good seed is pointed out by few good seeds. Thus, a complementary method is devised by considering photographers' visiting frequencies.

2) *Visiting Frequency Based Familiarity Score:* Since “familiarity” is time sensitive, on the basis of a v_i 's familiarity matrix T_i (see Definition 2) we calculate his/her “familiarity” of each $c \in C_i$ before a year λ . Three factors are considered:

- 1) For each year before λ , the proportion of days each c accounts for is recorded in T_i ;
- 2) Before λ , the reproducibility of cities is denoted as vector $\vec{\alpha}_i$. Assuming that before λ we collect 10 years' worth of image data of v_i in total, $\vec{\alpha}_i[c] = \frac{5}{10}$ if we detect that v_i went to c in five of the years;
- 3) The staleness of information. To identify whether v_i has been to these cities recently, we introduce a diagonal matrix, $\omega_i(x, y) \forall x, y \in \{1, 2, \dots, n\}$, to characterize this feature. n equals the number of detected years before λ .

Then, the score F_v is computed as follows.

$$F_v(v_i, c, \lambda) = \text{diag}(\vec{\alpha}_i^T \cdot (\omega_i \cdot T_i))[c] \quad (2)$$

where *diag* means matrix's diagonal vector, while $[c]$ is used to obtain the familiarity score of the target city c .

Obviously, the accuracy of F_v highly depends on the scale of the familiarity matrix T_i . For a photographer who has uploaded few images, this method would not be effective. To cover as many situations as possible, we need to discuss the fusion of F_s and F_v .

3) *Score Integration:* The most ideal situation is that F_v can be propagated over the network $G(V, E)$. Namely, we can merge these two scores by using F_v to select more good seeds (i.e., d_c in Eq. 1) in the social network based method. However, for the cities (e.g., Kyoto) where locals rarely upload images to Flickr, exceptions are raised:

- 1) F_s and F_v are almost independent. That means a photographer with high visiting frequency is not necessarily a resident. Inversely, a resident is very likely to have a low visiting frequency based score.
- 2) In accordance with the first issue, there is no similarity of visiting frequency between v_i and v_j , $\forall (v_i, v_j) \in E$. Namely, F_v cannot be used to select more good seeds.

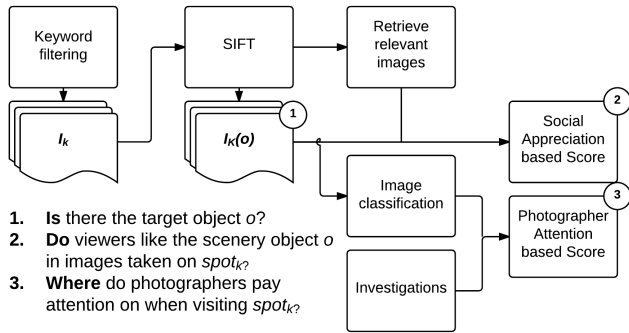


Fig. 3. Framework of ranking spots

For such exceptions that have no proportional relationships between F_s and F_v , we find that the best way of combining the two scores in our experiments is to take the maximum:

$$F(v_i, c, \lambda) = \max\{F_s(v_i, c), F_v(v_i, c, \lambda)\}. \quad (3)$$

B. Obscurity Score Calculation

For a target city c , we first introduce a *fuzzy set* $\{U_c, m(v_i) = F(v_i, c, \lambda)\}$ to specify each v_i 's familiarity on the basis of the integrated familiarity score $F(v_i, c, \lambda)$. Then, for a $spot_k$, we define two vectors, $\phi_f^{spot_k}$ and $\phi_{unf}^{spot_k}$, to represent the percentage of both the familiar group and unfamiliar group who have visited $spot_k$ in each year y :

$$\phi_f^{spot_k}[y] = \frac{w_y}{|U_c^k|} \sum_{v_i \in U_c^k} m(v_i); \quad \phi_{unf}^{spot_k}[y] = w_y - \phi_f^{spot_k}[y]. \quad (4)$$

Finally, to calculate the time sensitive familiarity before year λ , the entries in $\phi_f^{spot_k}$ and $\phi_{unf}^{spot_k}$ should include all the detected years before λ . We calculate $spot_k$'s obscurity score by comparing the norms of these two vectors.

$$ObscurityScore(spot_k, \lambda) = \left\| \phi_f^{spot_k} \right\| - \left\| \phi_{unf}^{spot_k} \right\|. \quad (5)$$

IV. SCENERY QUALITY EVALUATION

Figure 3 presents the framework of evaluating a $spot_k$'s scenery quality. By observing the behavior of both viewers (i.e., who have viewed the images in I_k) and photographers (i.e., who have taken the images in I_k), we defined three progressive criteria. Along with this flow chart, the three criteria will be introduced in the remainder of this section.

A. Image Labelling

Given a query consisting a city c and a scenery object o , the system labels the target images. Hereinafter, the set of all images about the scenery object o that are taken at $spot_k$ is denoted as $I_k(o)$. Benefiting from Flickr's well implemented keyword based search function, the labelling has already obtained a good performance. For example, through manually statistics, the percentage of noise images in $I_k(o = cherryblossom)$ is about 7.72 % on average.

We further applied SIFT [10] to reduce noise images in $I_k(o)$. In this model (see Figure 3), we use the key-points detected by SIFT to describe an image. For calculating the

similarity, a fast nearest-neighbor algorithm, FLANN, is used to do the matching by considering both numbers of matches and match distance. When filtering, $I_k(o)$ itself is used as a set of references. After mutual match, the system removes the images obtaining the lowest similarity scores. By sampling in our ranking experiments, the percentage of noise images finally dropped to about 5.86 %.

After noise image filtering, we have met the first and most elementary criteria that there is the scenery object o at $spot_k$.

B. Social Appreciation Based Analysis

In this method, we use an intuitive value, the number of "Favs", to evaluate a scenery object's quality in a social appreciation manner. "Fave this photo" is a common user behavior on SNS websites including Flickr for friends and visitors to mark on a message or an image. In [11], Pedro et al. used the number of "Favs" as ground truth for building and testing their image attractiveness ranking models. However, our task is faced with two new problems: (1) few images may be taken at $spot_k$; (2) a higher "Favs" may mean a beautiful image rather than beautiful scenery included in this image. In order to solve these two problems, two kinds of $I_k(o)$'s related images are introduced.

Similar Content Images: Similar to a semi-supervised learning method having a small training set, we expand the existing data sets (i.e., $I_k(o)$) by detecting the similar content images. For each image in $I_k(o)$, we retrieve its most similar images by applying SIFT model as shown in Figure 3. The newly expanded image set is denoted as $I_k^s(o)$. High social appreciation on $I_k^s(o)$ anyhow shows that the $spot_k$ is a good place for someone to take beautiful images about object o .

Context Images: We set the images from the same photosets to which an image $\eta \in I_k(o)$ belongs as its context images, denoted as C_η . Note that the image η itself is not in C_η . In this sense, both η and images in C_η have a similar chance of being accessed by viewers. In addition, all these images are likely taken by the same photographer as well as similar shooting skills. Under these contexts, compared with C_η , higher social appreciation on η suggests that η contains beautiful scenery.

On the basis of these two related image sets, the final social appreciation based score consists of two parts. The first $I_k^s(o)$ based part is computed by the mean of ratios χ between an image η 's number of "Favs" (i.e., $vote_\eta$) and view count (i.e., $view_\eta$):

$$S_1(spot_k, o) = \mu_\chi(I_k^s(o)); \quad \chi(\eta) = \frac{vote_\eta}{view_\eta}. \quad (6)$$

The second C_η based part is computed by z -scores:

$$S_2(spot_k, o) = \frac{1}{|I_k(o)|} \sum_{\eta \in I_k(o)} \frac{\chi(\eta) - \mu_\chi(C_\eta)}{\sigma_\chi(C_\eta)}. \quad (7)$$

Function $\mu_\chi(C_\eta)$ and $\sigma_\chi(C_\eta)$ are used to respectively calculate the average χ and χ 's standard deviation on the image set C_η . Thus, positive and negative z -scores represent observation

images above and below the mean, respectively. To combine S_1 and S_2 , a weight factor ω_c is used to adjust discrimination.

$$SAS(\text{spot}_k, o) = S_1 + \omega_c \cdot S_2. \quad (8)$$

By proposing the *SAS*, we have met the second criteria of answering where to take beautiful pictures. By adapting this system to be a photography suggestion guide, the value of ω_c is somewhat a trade-off of recommending either attractive images or beautiful scenery. However, nobody can indeed know the reasons a viewer “Faves” an image seen when browsing the Web. Sometimes, we “Fave” images just because our friends have posted them. By proposing a heuristic method in the next section, we try to evaluate the scenery quality from a photographer’s point of view.

C. Photographer Attention Based Analysis

For a scenery object o , it needs not only to be beautiful from a distance but also to have a delicate beauty up close. Figure 4(a.1) shows beautiful maple trees from far away, while in Figure 4(a.2) it shows a close-up of the beautiful leaves. The photographers’ choices of taking distance view or close-up view images psychologically indicate their attention to a scenery object o .

Accordingly, we first propose a new method to divide scenery images into “distance view” and “close-up view”, which exactly correspond to the two kinds of beauty above. Then, the regular patterns of photographers’ attention when visiting the known sightseeing spots with beautiful scenery object o are recognized. At last, we calculate the photographer attention based score, which indicates the degree of agreement with an observed regular pattern.

1) *Image Classification*: In the classification learning method, we use an image’s edge histogram and HSV (hue, saturation, and value) shown in Figures 4a and 4b as features to generate the training model. For the edge histogram, we find that a close-up view image has more clearer edges in the part where the photographer’s camera focused. In other words, the background of a close-up view image is more blurred. In contrast, a distance view image has a gentle contrast throughout the whole image. For the HSV, we assume that the hue and saturation values of close-up view images of particular scenery are always located in a relatively fixed range (see Figure 4b). In contrast, there is no such rule for distance view images. This is due to that, for close-up view images of particular scenery, most content of the images is generally made up of the target scenery objects.

In accordance with our experiments described in Figure 4, our classification model reaches a stable accuracy that is close to 80%. Torralba et al. [13] proposed a method to estimate the absolute depth of an image on the basis of the whole scene structure. For man-made scenes, the performance of depth estimations reaches 76%, while the result is relatively lower on natural image sets (about 70% on average). When applied in our task of classifying scenery images into two categories: (1) “distance view”; or (2) “close-up view”, our methods outperformed those of Torralba et al. [13].

2) *Photographer Attention based Scenery Score*: After the image classification, without loss of generality, we use the ratio of distance view images taken at a spot as r_o , which represents photographers’ attention on the scenery object o ’s beauty from a distance.

Figure 5a shows two observations on the known spots having beautiful maples and cherry blossom. For maples, we find that photographers paid more attention to the up-close beauty. Conversely, cherry blossom’s beauty from a distance attracted more attention compared with its up-close beauty.

Furthermore, by seeing the normal Q-Q plots in Figure 5b, we hypothesize that the two sample data sets come from the populations with normal distribution, since the points fall approximately along the 45-degree reference line. We performed the Lilliefors test, a kind of normality tests, to verify our hypothesis. By obtaining the p -value = 0.26 and 0.28, which are higher than the significance value (0.05 or 0.1), our hypothesis is statistically true.

As a result, based on the attention patterns recognized from spots having beautiful scenery o , a spot_k ’s photographer attention based score is computed using a probability density function. For instance, since maples and cherry blossom both have the same pattern, we use the normal distribution functions to get their scores:

$$PAS(\text{spot}_k, o) = \frac{1}{\sigma_o \sqrt{2\pi}} e^{-\frac{(r_o(k) - \mu_o)^2}{2\sigma_o^2}}. \quad (9)$$

μ_o and σ_o are obtained by distribution fitting of investigation results on the target scenery object o . $r_o(k)$ is the ratio of distance view images taken at spot_k . For different kinds of o , corresponding probability density functions are fitted. Intuitively, the *PAS* indicates how well a spot_k can attract visitors’ attention.

V. EXPERIMENTS

A. Data Preparation

From Flickr, we collected a total number of 929,403 images and the profiles of about 5,309,082 photographers. Using queries $q_1 = \text{“Kyoto, Maples”}$ and $q_2 = \text{“Kyoto, Cherry Blossom”}$, two target sub-datasets are labelled by our system. For the corresponding datasets, they contain 6,950 maple images taken by 177 photographers and 4,576 cherry blossom images taken by 319 other photographers, respectively.

B. Parameter Selection

Spot clustering method: Both density based clustering algorithm (e.g., [14]) and agglomerative methods like K-means can be used for building the graph H (see definition 3). Since the density based clustering would filter out a few sparsely distributed points that may be obscure, we utilized a centroid linkage based hierarchical clustering algorithm. We set the smallest cluster size to $2km$ (i.e., leaves of H), which enables us to find some obscure sightseeing spots that exist in famous sightseeing areas.

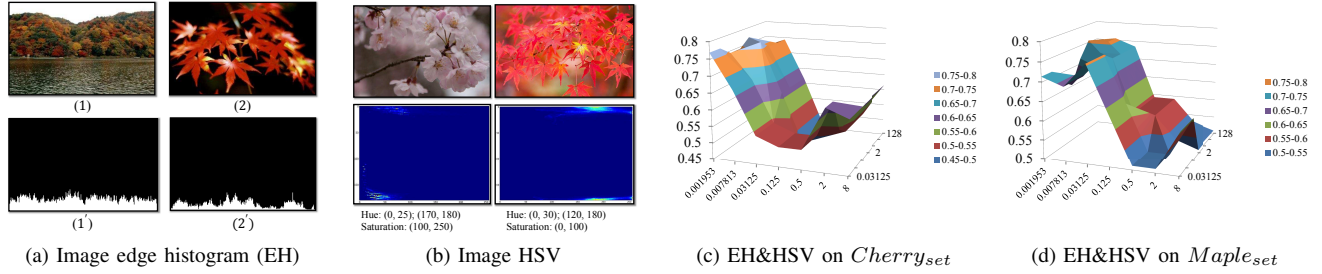


Fig. 4. Image classification. In (a), an image is vertically divided into 256 bins, and for each, the points of detected edges are counted to set up the histogram (i.e., EH). (b) presents the saturation (X-axis) and hue (Y-axis) distributions of close-up view images. Two original datasets, which respectively include 100 images, are labeled manually with either “distance view” or “close-up view”. Then, 3-fold cross validation is utilized. (c) and (d) illustrate the effectiveness of image classifiers on both the $Cherry_{set}$ and the $Maple_{set}$ by a grid search method proposed by Hsu et al. [12]. The X-axis ($\gamma = 2^{-9}, 2^{-7}, \dots, 2^3$) represents RBF kernel’s parameter, while the Y-axis ($C = 2^{-5}, 2^{-3}, \dots, 2^7$) stands for the penalty parameter in SVMs.

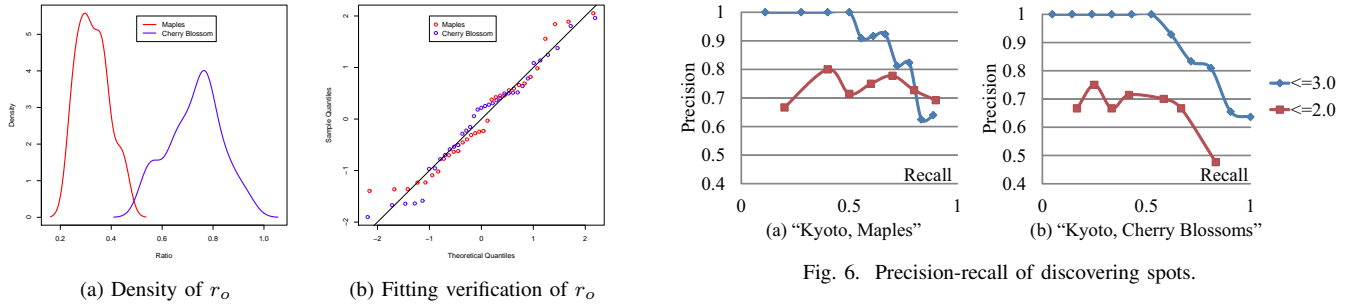


Fig. 5. Distribution fitting of sample data. Red: 22,421 geo-tagged maple images taken at 32 spots with beautiful maples; Blue: 10,748 geo-tagged cherry blossom images taken at 35 spots with beautiful cherry blossom.

TABLE I
DETAILS OF PHOTOGRAPHER CLASSIFICATION

| | |
|---|-------|
| Target photographers No. | 177 |
| Known Kyoto locals No. | 18 |
| Recall on social network based method | 50% |
| Recall on visiting frequency based method | 77.8% |
| Recall on using both of the two methods | 83.3% |

w : Since both w_i in Eq. 2 and w_y in Eq. 4 have the same meaning of characterizing the staleness of information, on the basis of the actual effectiveness, we set them to the same exponential decay: $\frac{1}{2}(\lambda^{-y})$. We set the w_c in Eq. 8 to 1.0, which means that the SAS emphasizes more on the beauty of scenery rather than the attractiveness of images.

C. Photographer Classification Evaluation

Table I shows the detailed information. Among the target 177 photographers, we obtain 18 Kyoto locals according to their public profiles, while the other 159 photographers are unknown. For the social network based method (i.e., Eq. 1), we retrieve 26,567 their friends to establish the social network. For the visiting frequency based method (i.e., Eq. 2), we totally collected 395,053 images taken by them to calculate the *familiarity matrices* T_i . The “recall” in Table I means the fraction of familiar photographers that are discovered among the 18 known residents.

For nodes (photographers) in the social network, we detect 318 residents as good seeds based on their available profiles. Although there are few good seeds selected, “familiarity communities” are still detected among the network. When we

set the threshold equal to 0.5 (i.e., $F_s(v_i, Kyoto) \geq 0.5$), the recall of this method reaches 50%. For the second method, on the basis of the normalized $F_v(v_i, Kyoto, 2013) \geq 0.5$, we obtain the recall of 77.8% among the 18 residents. The final recall obtained by considering both social network and visiting frequency (i.e., Eq. 3) is much higher, which reaches 83.3%.

D. Experiment on Obscurity Level Estimation

In order to analyze the effectiveness of our method for estimating a spot’s obscurity level (i.e., Eq. 5), we invited three residents who have lived in Kyoto for more than 20 years to label the discovered spots using an a five-point scale ranging from “0” for “not famous” to “4” for “very famous”.

Figures 6a and 6b illustrate precision-recall curves about the results of q_1 and q_2 . The blue curve shows the experimental results when we regard the spots with the average score of the three respondents with a relevance score no more than 3.0, while the red one displays the corresponding results when the relevance score is no more than 2.0.

By getting leaves of the tree-based graph H , we obtain 582 maple and 368 cherry blossom spots. To reduce the workload of respondents, we choose 32 maple candidate spots and 33 cherry blossom candidate spots from which 17 and 18 discovered obscure ones are returned, respectively, to do the analysis. The blue curves in both figures indicate that the *ObscurityScore* can exclude almost all the “very famous spots”. Furthermore, the red curves show that it still does well at discovering “not so famous” and “not famous” spots.

E. Experiment on Ranking Obscure Spots

Ground Truth: For q_1 and q_2 , we obtained the ground truth through field visits conducted in November 2013 and April

TABLE II
TWO CRITERIA OF SCENERY RANKING SCORES

| Criteria | Explanations | Ratings |
|----------|--|---------|
| v | Is the spot worth visiting or not. | 0,1,2,3 |
| r | Is the spot worth recommending or not. | 0,1,2,3 |

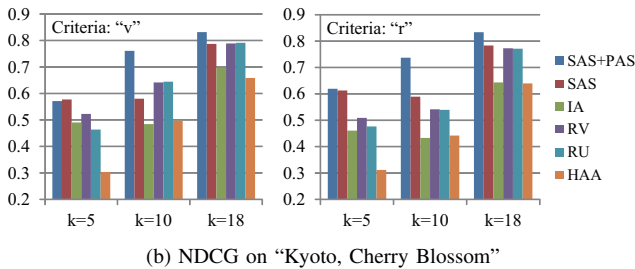
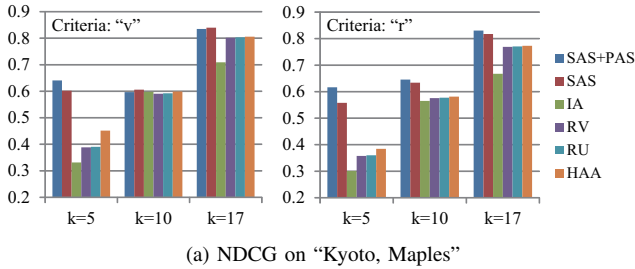


Fig. 7. NDCG values of the top- k ranking results.

2014. For each of the selected 32 maple spots and 33 cherry blossom spots, we invited at least three subjects to perform field evaluation. Using Google Nexus 7, we asked them to record their visiting experiences on Twitter as well as take geo-tagged photos at each spot. Verified by these logs, we say that the human evaluation is credible. After that, they answered a questionnaire about their opinions of the spots they had visited. On the basis of the feedback, we defined two criteria to denote how well a discovered obscure spot meets tourist requirements shown in Table II.

We compare other 4 baseline methods with the SAS (i.e., Eq. 8) and PAS (i.e., Eq. 9):

- Image Attractiveness (IA): the image attractiveness ranking model proposed in [11];
- HITS based Authority Analysis (HAA): the user(hub)-location(authority) graph based analysis in [1];
- Rank-by-Visits (RV): the accumulative counts of visits to the spot;
- Rank-by-Users (RU): the accumulative counts of unique users who have been to the spot.

By implementing these methods, we rank the 17+18 obscure spots mentioned above.

NDCG based Evaluation: Figure 7a and 7b, respectively, depict the NDCG [15] of different measures on the two queries: q_1 and q_2 . From the data shown in these figures, we can obtain the following results.

First, in contrast to the baseline methods, our proposed methods, $SAS + PAS$ and SAS , performed more effectively in the scenery quality based obscure spot ranking.

Second, the PAS brings a significant improvement to the ranking only based on SAS . Since ranking obscure spots of having few images is very challenging, our experiments verified that the photographer attention based analysis could provide a stable effectiveness improvement.

Third, by comparing Figure 7a and 7b, the authority based analysis, HAA , fluctuated considerably. Namely, this series of conventional methods heavily depend on how much the corresponding raw data obtained.

Finally, the results related to IA have verified our hypothesis that an attractive image doesn't mean beautiful scenery included in this image. As naive methods, RV and RU both performed better than expected, because our system has pre-filtered out the noises (like crowded spots of having no target scenery) in the first step of ranking.

VI. CONCLUSION

We proposed an obscure sightseeing spots discovery system based on two novel metrics, obscurity level and scenery quality. In the experiments, we employed human effort to obtain ground truth for the two query instances and further analyzed the performance of each proposed method by comparison.

ACKNOWLEDGMENT

This work is supported by JSPS KAKENHI Grant Numbers 25700033, 15J01402, and 15K00236.

REFERENCES

- [1] Y. Zheng and X. Xie, "Learning travel recommendations from user-generated gps traces," *ACM TIST*, vol. 2, no. 1, p. 2, 2011.
- [2] W.-C. Chen, A. Battestini, N. Gelfand, and V. Setlur, "Visual summaries of popular landmarks from community photo collections," in *ACM Multimedia*, 2009, pp. 789–792.
- [3] J. Liu, Z. Huang, L. Chen, H. T. Shen, and Z. Yan, "Discovering areas of interest with geo-tagged images and check-ins," in *ACM Multimedia*, 2012, pp. 589–598.
- [4] K. Hasegawa, Q. Ma, and M. Yoshikawa, "Trip tweets search by considering spatio-temporal continuity of user behavior," in *DEXA (2)*, 2012, pp. 141–155.
- [5] Y. Zheng, Z. Zha, and T. Chua, "Research and applications on georeferenced multimedia: a survey," *Multimedia Tools Appl.*, vol. 51, no. 1, pp. 77–98, 2011.
- [6] C. Zhuang, Q. Ma, X. Liang, and M. Yoshikawa, "Anaba: An obscure sightseeing spots discovering system," in *ICME*, 2014, pp. 1–6.
- [7] Flickr apis. [Online]. Available: <https://www.flickr.com/services/api/>
- [8] Z. Gyöngyi, H. Garcia-Molina, and J. O. Pedersen, "Combating web spam with trustrank," in *VLDB*, 2004, pp. 576–587.
- [9] D. Xu, P. Cui, W. Zhu, and S. Yang, "Find you from your friends: Graph-based residence location prediction for users in social media," in *ICME*, 2014, pp. 1–6.
- [10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [11] J. S. Pedro and S. Siersdorfer, "Ranking and classifying attractiveness of photos in folksonomies," in *WWW*, 2009, pp. 771–780.
- [12] C. wei Hsu, C. chung Chang, and C. jen Lin, "A practical guide to support vector classification," 2010.
- [13] A. Torralba and A. Oliva, "Depth estimation from image structure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1226–1238, 2002.
- [14] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *KDD*, 1996, pp. 226–231.
- [15] K. Järvelin and J. Kekäläinen, "Cumulated gain-based evaluation of ir techniques," *ACM Trans. Inf. Syst.*, vol. 20, no. 4, pp. 422–446, 2002.