

# Chapter 10

## Personalized Travel Navigation and Photo-Shooting Navigation Using Large-Scale Geotags

T. Yamasaki, A. Gallagher, T. Chen and K. Aizawa

**Abstract** In this chapter, a geotag-based inter/intra-city travel and photo-shooting navigation system that considers both the personal preference and the seasonal/temporal popularity is presented. For the inter-city travel navigation, similarity among users is efficiently calculated by combining our visit pattern similarity and photo shooting pattern similarity. Accurate intra-city travel navigation is achieved by incorporating the seasonal and temporal information into a Markov model. Photo-shooting navigation by using large-scale geotagged photos is also presented. The effectiveness of the proposed algorithms have been experimentally demonstrated by using more than millions of geo-tags and photos downloaded from Flickr.

### 10.1 Introduction

Travel planning is fun but sometimes troublesome especially when the travelers are not familiar with the city. Therefore, we often refer to travel guidebooks such as Verlag Karl Baedeker,<sup>1</sup> the Michelin Guide,<sup>2</sup> and Lonely Planet.<sup>3</sup>

---

<sup>1</sup> <http://www.baedeker.com>.

<sup>2</sup> <http://www.michelintravel.com>.

<sup>3</sup> <http://www.lonelyplanet.com>.

T. Yamasaki (✉) · K. Aizawa  
Department of Information and Communication Engineering,  
The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan  
e-mail: yamasaki@hal.t.u-tokyo.ac.jp

K. Aizawa  
e-mail: aizawa@hal.t.u-tokyo.ac.jp

A. Gallagher · T. Chen  
School of Electrical and Computer Engineering, Cornell University,  
Ithaca, NY 14853-7501, USA  
e-mail: andrew.c.gallagher@gmail.com

T. Chen  
e-mail: tsuhan@ornell.edu

In the Internet era, we can also check user-generated travel assistance WEB sites such as Yelp,<sup>4</sup> TripAdvisor<sup>5</sup> and Yahoo! Travel.<sup>6</sup> However, most of the travel guides mentioned above only provide a ranking list of the landmarks. Namely, they do not provide how (in what order) we should travel around those landmarks and they do not consider the travelers' personal preference, either. In addition, it is often difficult to know detailed travel information such as what the best season or the best time of the day is to visit a certain landmark.

In recent years, travel recommendation and navigation using large-scale user-generated photos has been a hot topic because such photos contain rich meta data such as tags, time, and geo-locations (or geo-tags). Although there are also a lot of photos without such meta data, location of the photos could be estimated using the location estimation techniques in [9, 13, 15, 24, 32, 34] and those introduced in this book.

In this chapter, we propose both travel and photo-shooting navigation based on statistical analysis of such large scale user-generated photos and accompanying meta data. As a result, we can navigate users considering the personal preference and seasonal/temporal information. For the inter-city navigation, collaborative filtering based similar traveler extraction is presented. The similarity among the travelers is calculated by considering several user behavior patterns such as visiting patterns and photo-shooting patterns. For the intra-city travel navigation, the Markov-based behavior model considering seasonal/temporal information is demonstrated. Regarding the photo-shooting navigation, "where to go" and "how to shoot" is recommended based on the location-specified similar image retrieval.

The experiments for the travel navigation were conducted using 6.2 million geo-tag data from 21 famous cities and parks in the world collected from Flickr.<sup>7</sup> Experimental results show that the proposed model drastically outperforms the previous travel navigation systems. In addition, we demonstrate some photo-shooting navigation examples using 2.2 million geo-tags and photos in three cities.

The rest of this chapter is organized as follows. Section 10.2 reviews recent related works. The algorithms for our navigation systems are summarized in Sects. 10.3 and 10.4. The experimental results are presented in Sect. 10.5, followed by concluding remarks in Sect. 10.6.

## 10.2 Related Works

### 10.2.1 Travel Navigation

GPS-embedded cell phones and cameras and the popularity of photo sharing sites have enabled us to analyze how people travel around the globe. Geotagged

---

<sup>4</sup> <http://www.yelp.com>.

<sup>5</sup> <http://www.tripadvisor.com>.

<sup>6</sup> <http://travel.yahoo.com>.

<sup>7</sup> [www.flickr.com](http://www.flickr.com).

photos are the rich source of information because they contain not only spatial and temporal information of the users but also text tags. It is also possible to analyze the preference of the users by analyzing the number of pictures taken at certain locations. In other words, the user generated geo-tag information can be regarded as human sensors [14]. For instance, it has been demonstrated that Italian travelers and American travelers tend to take different travel routes in Rome [14]. This investigation shows the potential of analysis of the large-scale geotagged data.

By analyzing the geo-tags and the text tags, it is possible to automatically extract landmarks of the city [7, 12, 21, 31, 38], rank the landmarks [18], detect particular events [21, 30], detect object of interest [30], extract representative photos of the landmarks [12, 20, 21], detect panoramic view spots [29], and so on.

Travel route recommendation is one of such applications. Ying et al. [37] recommended travel routes depending on how many hours users can spend in the landmark or the city. [10] summarized the geo-tags of Flickr photos and solved the intra-city travel recommendation as a orienteering problem. Kurashima et al. [22] proposed a probabilistic behavior model by combining topic models and Markov models. Personal preference was implicitly included in the topic model. The system could also suggest travel routes considering user-specified time periods. Arase et al. [2] categorized the geotagged photos into six trip patterns such as landmark visiting, enjoying nature, daily trips in local area, and so on. Users could browse typical photos of the six categories to decide where to go. Cheng et al. [8] analyzed faces in the photos and showed that the travelers' attributes such as gender, age, and race can facilitate better travel route recommendation. Such techniques can also be combined with automatic travel guide generation systems [6] and drive navigation systems considering the scenic attractiveness [39].

There are only a few approaches on inter-city travel recommendation as far as we know [11, 19, 26, 28]. The inter-city recommendation is to recommend places to visit in the city A to those who have traveled in the city B (but never been to the city A). [11, 26] analyzed travelers' similarity by collaborative filtering. Popescu et al. [28] calculated the travel pattern similarity between users based on the kernel convolution using the raw geo-tag data. Jiang et al. [19] introduced more features such as similarity of the tags and visual similarity of the photos between users.

### ***10.2.2 Photo-Shooting Navigation***

Obtaining good photos as a memory of the travel has been an important but at the same time difficult problem for us. In fact, we can find a lot of photo-shooting guidebooks in bookstores.

Attractiveness enhancement of photos is one of the solutions and has been one of the main topics of image processing and computer graphics such energy-based inpainting/retouching [3, 4, 33], and compositing [1]. Example-based processing using large-scale photos on the Internet [16] is another hot topic in image manipulation. Photo quality assessment [5, 36] and photo editing based on the quality measure [23, 25, 27] can also be used for obtaining good photos.

However, these approaches can be used for post-processing photos, only after coming back from the travel. On the other hand, navigating “where to go” and “how to take” has not been investigated deeply so far.

### 10.2.3 Contribution of this Work

The contribution of this work is as follows: (1) For the inter-city recommendation, similarity between users is calculated by considering visit pattern similarity, photo shooting pattern similarity and photo preference pattern similarity. In addition, seasonal information is considered. (2) For the intra-city recommendation, seasonal and temporal information is considered and incorporated into the Markov model for better travel recommendation. (3) Photo-shooting navigation such as where to go and how to shoot in order to obtain good photos is addressed by using large-scale geotagged photos on the Internet. Note that this chapter is the extended version of [35].

## 10.3 Travel Navigation

### 10.3.1 Inter-city Travel Navigation

The purpose of the inter-city recommendation is to use the user’s previous travel patterns in certain cities in order to recommend “must-see” landmarks when the user is visiting another city. The assumption is that the travelers who had similar travel patterns with the current user in a certain city would also have similar travel patterns in another city. The user’s previous travel patterns are useful for the inter-city recommendation because the travel patterns implicitly indicates the user’s preference. In this chapter, collaborative filtering based approach is used. Here, let us assume that the user has visited the city  $c$  previously and wants to visit the city  $d$ . In the collaborative filter, travelers who have traveled both in the cities  $c$  and  $d$  are collected. And those who had similar travel patterns in the city  $c$  with the current user are extracted. Then, the landmarks in the city  $d$  are re-ranked by their travel patterns. Therefore, how to calculate the similarity between users is a key issue. In this chapter, two similarity measures between users are defined as follows.

#### Visit pattern similarity

The visit pattern similarity considers whether the user has visited certain landmarks only. Therefore, the vector for the user  $i$  in the city  $c$  is defined as:

$$\mathbf{v}_i^c = (v_{i1}^c, v_{i2}^c, \dots, v_{iL^c}^c) \quad (10.1)$$

where  $L^c$  is the number of landmarks in the city  $c$ .  $v_{il}^c = 1$  if the user has visited the landmark  $l$  and  $v_{il}^c = 0$  if not. Then, the similarity between the user  $i$  and the user  $j$

is defined as the cosine similarity:

$$\text{Sim}_v^c(\mathbf{v}_i^c, \mathbf{v}_j^c) = \frac{\mathbf{v}_i^c \cdot \mathbf{v}_j^c}{|\mathbf{v}_i^c| |\mathbf{v}_j^c|}. \quad (10.2)$$

### Photo shooting pattern similarity

The photo shooting pattern considers how the user liked the landmarks, which is modeled as a function of the number of pictures taken at the landmarks:

$$\mathbf{p}_i^c = (p_{i1}^c, p_{i2}^c, \dots, p_{iL^c}^c) \quad (10.3)$$

$$p_{il}^c = \log(N_{il}^c + 1) \quad (10.4)$$

where  $N_{il}^c$  is the number of photos at the landmark  $l$  in the city  $c$  taken by the user  $i$ . The similarity is calculated as:

$$\text{Sim}_p^c(\mathbf{p}_i^c, \mathbf{p}_j^c) = \frac{\mathbf{p}_i^c \cdot \mathbf{p}_j^c}{|\mathbf{p}_i^c| |\mathbf{p}_j^c|}. \quad (10.5)$$

The total similarity between the user  $i$  and  $j$  is calculated by the weighted sum of the visit pattern similarity and the photo shooting pattern similarity:

$$\text{Sim}_{\text{total}}^c(u_i^c, u_j^c) = \alpha \text{Sim}_v^c(\mathbf{v}_i^c, \mathbf{v}_j^c) + (1 - \alpha) \text{Sim}_p^c(\mathbf{p}_i^c, \mathbf{p}_j^c) \quad (10.6)$$

where  $u_i^c$  represents the user  $i$  in the city  $c$ .

In this chapter, the travelers whose total similarities are greater than a certain threshold  $\text{Sim}_{th}$  are regarded as similar travelers. After extracting the similar travelers by Eq. 10.6 from the city  $c$ , the score for each landmark in the city  $d$  is calculated as follows:

$$\text{Score}_{k,s}^d = \sum_m \left( \alpha \mathbf{v}_{m,s}^d + (1 - \alpha) \mathbf{p}_{m,s}^d \right) \quad (10.7)$$

where  $\text{Score}_{k,s}^d$  represents the score for the  $k$ th landmark in the city  $d$ , which considers the seasonal popularity.  $\mathbf{v}_{m,s}^d$  and  $\mathbf{p}_{m,s}^d$  represent the visit pattern score and the preference pattern score of the  $m$ th user in the city  $d$  in the season  $s$ , respectively. Here, the users are those who had the similarity greater than  $\text{Sim}_{th}$  with the user  $i$  in the city  $c$  as discussed above. Then, the ranking of the landmarks in the city  $d$  is calculated based on Eq. 10.7. Note that the score for the landmarks in Eq. 10.7 is dependent on the user and the season. Therefore, the landmarks to visit in the city  $d$  are recommended by analyzing similar travelers' travel pattern and seasonal attractiveness.

Different from intra-city recommendation, the temporal information is not considered because the inter-city recommendation is for generating a list of landmarks

to visit, not the order of landmarks to visit. Once the user actually starts traveling in the city  $d$ , the route can be recommended by the intra-city recommendation.

### 10.3.2 Intra-city Travel Navigation

Let us assume that we are traveling in NYC. For instance, Rockefeller center is one of the most popular landmarks in NYC, but it becomes particularly popular at night in winter. Because people want to take photos of the Christmas tree, the statue of Prometheus with the ice skate link, and clear night view from the top of the rock. It is also recommended to visit the statue of liberty before Rockefeller center because the statue of liberty is closed at night. These are just intuitive examples of how our algorithm can help users to find the most appealing landmark in the city. Small festivals or events are not usually introduced in travel guidebooks. Some famous buildings might be closed at night but it might be possible to take pictures from outside. Therefore, a travel recommendation system that can automatically extract temporal/seasonal popularity is required to handle these issues. And also, the system can give us an important information when to visit to enjoy the landmark the best.

Similar to Cheng’s model [8], we generate a touring model using a Markov model. Instead of using users’ profile such as gender and race as in [8], we introduce seasonal and temporal information ( $\mathbf{S}_u$ ).

$$L^* = \arg \max_{L_j} P(L_j | \mathbf{S}_u, L_i) \quad (10.8)$$

$$\mathbf{S}_u \in (s, t) \quad (10.9)$$

where  $s$  is the season and  $t$  is the time of the day. Namely, the next destination is dependent only on the current user’s location and the other landmarks the user previously has visited do not matter. This model is practical because it can consider the “temporal distance”. For instance, even when the landmarks A and B are geometrically far away, there might be a public transportation and would take only a few minutes. In such a case, the transition probability from the landmark A to the landmark B would become high. Although the Markov model can suggest only the next place to visit, it can be used for travel route recommendation because we can repeat the recommendation by eliminating already visited landmarks. In fact, the Markov model-based approaches are often used for the travel recommendation as in [8, 22]

The simplest approach is to separate the geotagged data into different seasons and time ranges and generate the transition probability matrices from the current location  $L_i$  to the next location  $L_j$ . This transition model works fine if the number of travelers in the city is large enough such as in NYC. On the other hand, in smaller cities, the transition matrix tends to be sparse and might include only a few (or no) travelers in a particular season and time. For instance, when we divide the data into four seasons and four time spans (i.e., dividing a day into 6 h bins: morning, afternoon, night, late at night), the travelers would be divided into 16 different transition matrices, which would prevent proper recommendation.

Another solution is to use the Bayes' theorem as in [8]. The probability that the location  $L_j$  to be recommended when the user  $u$  with the seasonal and temporal information  $\mathbf{S}_u$  is at the location  $L_i$  is described as:

$$P(L_{i \rightarrow j} | \mathbf{S}_u) = \frac{P(L_{i \rightarrow j}, \mathbf{S}_u)}{P(\mathbf{S}_u)} \quad (10.10)$$

$$= \frac{P(L_{i \rightarrow j})P(\mathbf{S}_u | L_{i \rightarrow j})}{P(\mathbf{S}_u)} \quad (10.11)$$

$$= \frac{P(L_i)P(L_j | L_i)P(\mathbf{S}_u | L_{i \rightarrow j})}{P(\mathbf{S}_u)} \quad (10.12)$$

Therefore, the Eq. (10.8) will become

$$L^* = \arg \max_{L_j} P(L_j | L_i)P(\mathbf{S}_u | L_{i \rightarrow j}) \quad (10.13)$$

because  $P(L_i)$  and  $P(\mathbf{S}_u)$  are independent of  $L_j$ . If we assume the independence between  $s$  and  $t$ , the joint probability  $P(\mathbf{S}_u | L_{i \rightarrow j})$  can be rewritten as:

$$L^* = \arg \max_{L_j} P(L_j | L_i)P(s | L_{i \rightarrow j})P(t | L_{i \rightarrow j}) \quad (10.14)$$

$P(L_j | L_i)$ ,  $P(s | L_{i \rightarrow j})$ , and  $P(t | L_{i \rightarrow j})$  can be estimated from the training data:

$$P(L_j | L_i) = \frac{\text{count}(L_{i \rightarrow j})}{\sum_{j \in \mathbf{L}} \text{count}(L_{i \rightarrow j})} \quad (10.15)$$

$$P(s \text{ or } t | L_{i \rightarrow j}) = \frac{\text{count}(L_{i \rightarrow j} \cap \mathbf{S}_u = s \text{ or } t)}{\text{count}(L_{i \rightarrow j})} \quad (10.16)$$

where  $\text{count}(L_{i \rightarrow j})$  the total number of travelers who traveled from the location  $L_i$  to  $L_j$ , and  $\text{count}(L_{i \rightarrow j} \cap \mathbf{S}_u = s \text{ or } t)$  is that in a specific season and time of the day.

The model we propose in this chapter is a naive transition model:

$$L^* = \arg \max_{L_j} P(L_j | L_i)P(s | L_{i \rightarrow j})P(t | L_{i \rightarrow j}) \quad (10.17)$$

Therefore, the the transition model is divided into a general model, seasonal transition model, and a temporal transition model.

Note that our seasonal and temporal modeling is orthogonal to the previous Markov model-based approaches [8, 22] and thus can be integrated into them, as well.

## 10.4 Photo-Shooting Navigation

The travel guides usually tell us what we can see when we visit the landmarks. On the other hand, when we want to take photos of landscape of the city, for instance, we need to look for such landmarks by reviewing all the landmarks in the guidebook.

In addition, taking “good” photos is another difficult problem for the travelers. Therefore, we propose a photo-shooting recommendation system that can navigate the users where to go and how to take in order to shoot photos they really want to take.

Assume that the user is now in the city  $c$  has a landscape photo which was either taken by the user himself/herself or retrieved from the Internet by the keyword search. Here, we do not care about where the landscape photo was taken (It might have been taken in the city  $c$  or in another city). Then, we extract the GIST descriptor from the landscape image. GIST descriptors are extracted for the photos taken in the city  $c$ , as well. Since the GIST descriptor is good at retrieving similar scene/composition images, we can retrieve similar landscape images taken in the city  $c$  and show the locations of the retrieved photos on the map. This scenario is applied not only to landscape photos, but also to buildings, streets, crowds, nature, and so on. In this manner, the system can navigate where to go to take photos of a certain taste.

The example above is photo-to-photo-similarity-based recommendation. Similarly, user-to-user similarity-based recommendation can also be achieved. When the user inputs several photos they like, regardless of where they were taken, we can retrieve the users who have taken similar photos in the city  $c$  and visualize where such similar users visit and what kind of pictures they take.

It sometimes happens that the photos we have taken at landmarks are not as good as those on the guidebooks. Therefore, we also present a “how to shoot” navigation system using large-scale photos on the Internet. Once the user’s location is identified, the photos taken around that area are retrieved and sorted by the order of photo attractiveness. To evaluate such attractiveness, we use a model in [17]. Reference [17] demonstrated that the attractiveness of the photos can be roughly estimated as a function of the number of views, the number of favorites, and the number of days the photo is revealed on the Internet as follows:

$$(\text{Attractiveness score}) = \frac{(\# \text{ of favorites})^2}{(\# \text{ of views})} + \gamma((\# \text{ of views}) - m) \quad (10.18)$$

$$m = (\text{average number of views of the photos uploaded on the same day}) \quad (10.19)$$

The advantage of this model is that the attractiveness score can be estimated only from the meta data. When the users want to use more accurate attractiveness evaluation, image-based approaches introduced in 10.2.2 can be used.

## 10.5 Experimental Results

### 10.5.1 Experimental Setup

The datasets used in the travel navigation were collected by crawling Flickr using its public API by ourselves. The cities crawled were 21 cities and parks in the world (Boston, Chicago, New York, Philadelphia, San Diego, San Francisco, Seattle, Toronto, Washington D.C., Yosemite, Yellowstone, Niagara, Honolulu, Las Vegas,



Taipei, Dallas, Praha, Kyoto, Vancouver, Firenze, Brisbane). All photos were taken between January 1, 2004 and June 31, 2011. The geo-tag information was clustered by using mean shift by following previous papers [12, 22] to find landmarks effectively. The bandwidth was set as 100m. The landmarks with less than 100 unique users were eliminated. As a result, the data consists of 6,253,865 photographs and their associated meta data, which were taken by 219,390 unique users.

We evaluated the performance by using all the landmarks, the top 25 landmarks, and the top 10 landmarks. For instance, if the user traveled

$$L_1^5 \rightarrow L_2^{24} \rightarrow L_3^1 \rightarrow \dots, \quad (10.20)$$

and if we look at only the top 10 landmarks, the travel route is regarded as

$$L_1^5 \rightarrow L_2^1 \rightarrow \dots, \quad (10.21)$$

where  $L_i^k$  means the popularity of the  $i$ th location is ranked at the  $k$ th place. When a user is traveling from the landmark  $A$  to the landmark  $B$ , the user tends to visit other landmarks on the way to  $B$  even if they are not very popular (i.e., Charging Bull on the way from Ground Zero to Battery park in New York). This process neglects such minor landmarks.

The recommendation was evaluated by the leave-one-user-out method. Namely, the recommendation model was generated for each user using all the other travelers' history in each city.

For photo-shooting navigation, we have collected 2.2 million geotagged photos from Flickr, which were taken in New York, Paris, and Tokyo. The photos were taken from January 1, 2008 to December 31, 2012. Since it is difficult to evaluate the photo-shooting recommendation accuracy, only the results are shown below.

### 10.5.2 Inter-city Travel Navigation

For the inter-city recommendation, the users who have traveled two or more cities (out of the 21 cities listed above) and have visited two or more landmarks in each city were extracted. In addition, the city/park pairs that contained at least 500 travelers were considered. As a result, 17,016 users and 36 city/park pairs were detected.

The accuracy was evaluated by the mean average precision (mAP) of the recommended landmarks:

$$mAP = \frac{1}{Q} \sum_{q=1}^Q AP_q \quad (10.22)$$

$$AP_q = \frac{1}{P_q} \sum_{p=1}^{P_q} \frac{p}{(\text{the order of } p\text{th hit})} \quad (10.23)$$

where  $Q$  is the number of travelers,  $AP_q$  is the average precision of the  $q$ th traveler, and  $P_q$  is the number of landmarks the traveler  $q$  visited. Namely, the order of landmarks to visit was not considered and only the ranking of the recommended landmarks are evaluated.  $\text{Sim}_{th}$  and  $\alpha$  were set to 0.1 and 0.6, respectively, in this experiment.

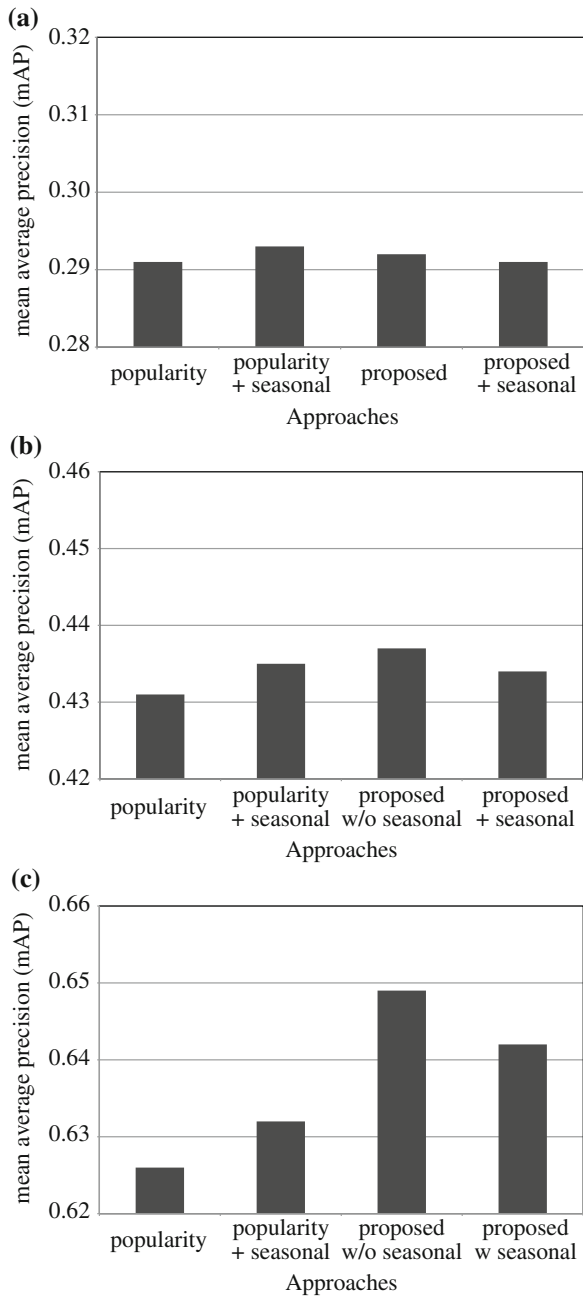
The mean average precision values are shown in Fig. 10.1. Our proposed method is compared with two baselines: the simple popularity-based recommendation, which is generated solely from the popularity ranking, and seasonal-popularity-based recommendation, which is based on season-aware landmark popularity ranking. It is demonstrated that our proposed inter-city recommendation model is better than the simple popularity-based recommendation and seasonal-popularity-based recommendation when top 10/25 landmarks were considered. For instance, mAP is better by more than 0.02 when the top 10 landmarks are considered. It is observed that the proposed model with the seasonal information is worse than that without the seasonal information. It is because we do not have enough number of travelers in constructing the model. It is expected that the proposed model with the seasonal information would be improved if more travelers' data are accumulated. On the other hand, our proposed model gets worse than the baseline methods when we consider all the landmarks. This is also because of the lack of the number of travelers. When we increase the dimension of the similarity pattern vector (the number of landmarks), the similarity score would become sensitive to noise especially when the number of travelers is small.

The impact of changing  $\text{Sim}_{th}$  and  $\alpha$  is demonstrated in Fig. 10.2.  $\text{Sim}_{th}$  should be 0–0.2, which indicates that the similarity-based thresholding does not contribute to better recommendation. However, we think that this is because the number of similar travelers becomes too small when we set  $\text{Sim}_{th}$  larger. It is advised to set  $\alpha$  as 0.4–0.8, showing that visit pattern similarity is more informative than photo shooting pattern similarity.

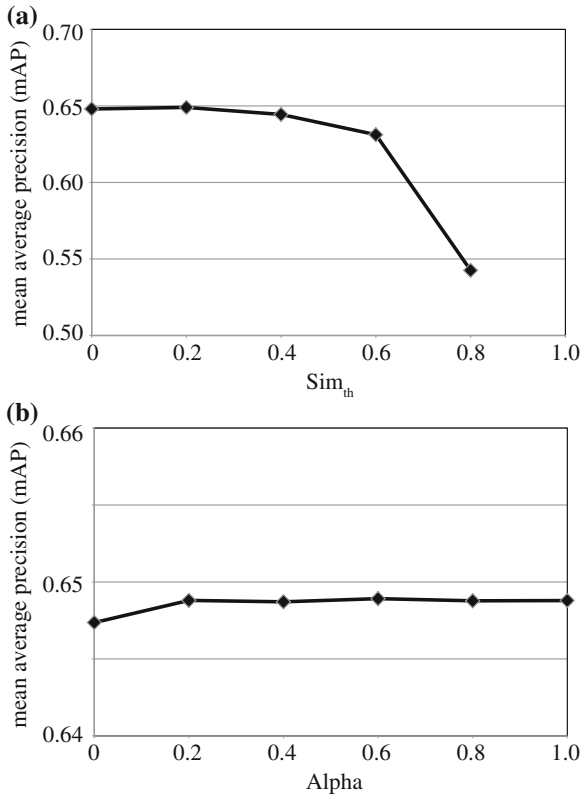
### 10.5.3 Intra-city Travel Navigation

For the intra-city recommendation, all the 6,253,865 geo-tags by 219,390 unique users were used. The performance was evaluated by measuring how accurately the system estimated the next location  $L_j$  when the user was at the location  $L_i$ . As described in Sect. 10.5.1, the recommendation accuracy was calculated by the leave-one-user-out method. Namely, one user is used as test data and the other users' travel history were used to build the model. In the experiment, we compared our approach with the following probabilistic models:

- Multinomial model: predicts the next landmark based on its popularity. The most popular landmark except for the current and already visited ones is recommended. This model does not consider the user's current location.
- Markov model: recommend the next landmark based on the user's current location using  $P(L_j|L_i)$ . This model considers the user's current location but does not consider the seasonal or temporal information.



**Fig. 10.1** Mean average precision of the inter-city recommendation: **a** all the landmarks, **b** top 25 landmarks, and **c** top 10 landmarks



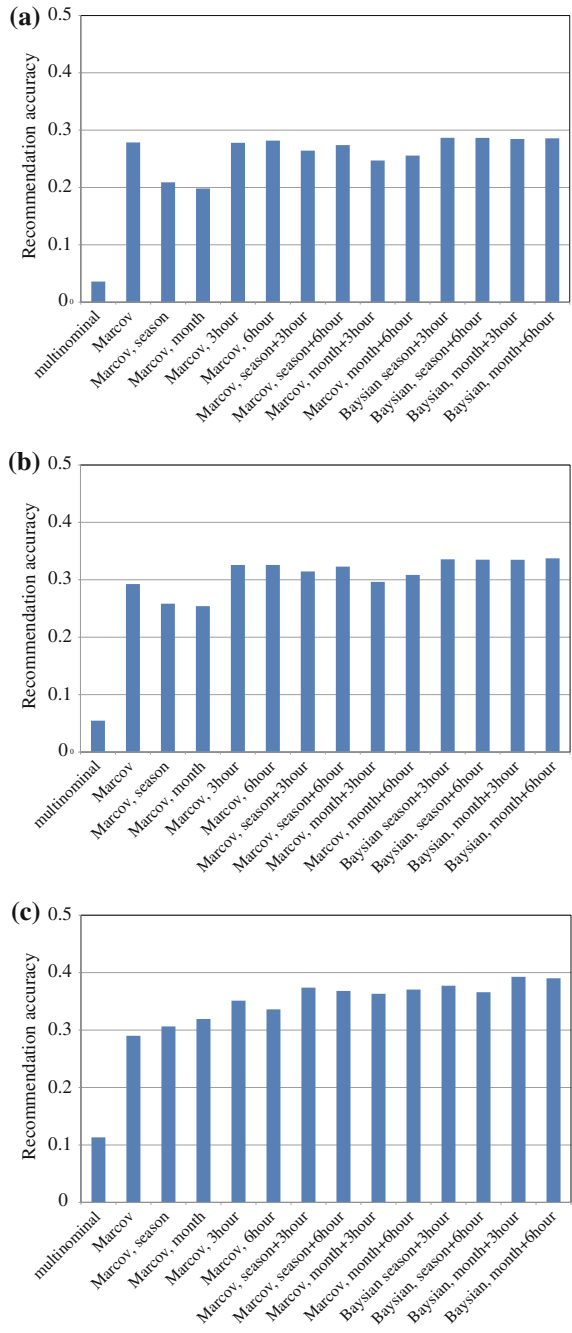
**Fig. 10.2** Mean average precision as a function of **a**  $Sim_{th}$  **b**  $\alpha$ . Top 10 landmarks were considered

- Seasonal/Temporal Markov model: recommend the next landmark based on the user's current location using  $P(L_j, (s \text{ or } t)|L_i)$ . This joint model that considers the user's current location and either of the seasonal or temporal information.
- Seasonal&Temporal Markov model: recommend the next landmark based on the user's current location using  $P(L_j, s, t|L_i)$ . This joint model that considers the user's current location, the seasonal information and the temporal information.

For each seasonal, temporal, and seasonal & temporal (proposed) model, two different seasonal information (4 seasons and 12 months) and two different time divisions (every 6 h and every 3 h) were trained and tested.

The detailed analysis of the recommendation accuracy is demonstrated in Fig. 10.3. Although slight difference can be observed from each other, the recommendation model considering the seasonal and temporal information always works better than the original Markov model. It is also demonstrated that the temporal information contributes more in the travel navigation when we look at the third to sixth columns in the figures. In addition, our Bayesian-based model is always better than the joint model in which the geotagged data are separated into different seasons and time

**Fig. 10.3** Detailed mean accuracy of the next location recommendation: **a** all landmarks, **b** top 25 landmarks, and **c** top 10 landmarks



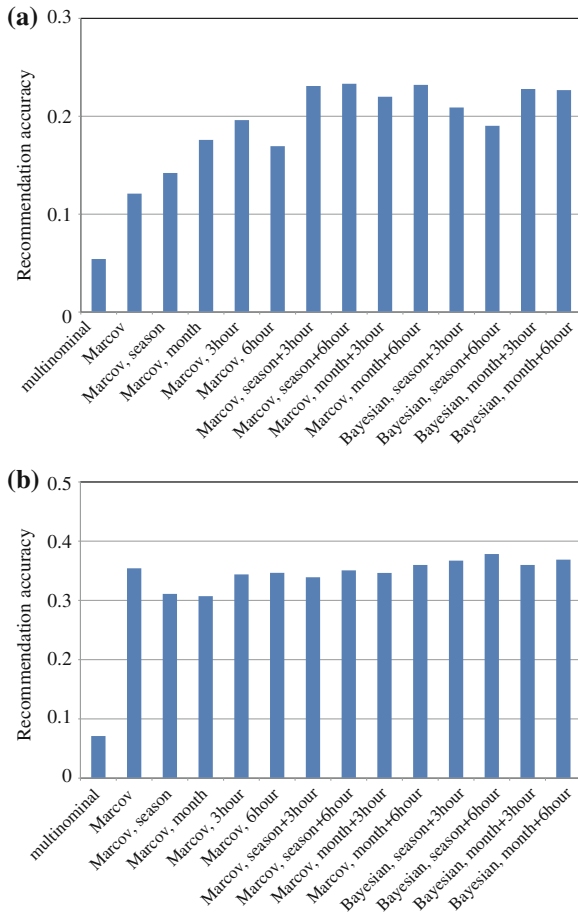
**Table 10.1** Accuracy of intra-city recommendation

	Naive Markov (%)	Their results (%)	Improvement from Markov (%)
Eastern cities in US [22] (bw = 50m)	35.4	37.0	+1.6
Western cities in US [22] (bw = 50m)	24.7	27.2	+2.5
New York [8] ( $H(L_{ij}) \geq 2$ )	28	28	+0.03
	Naive Markov (%)	Our results (%)	Improvement from Markov (%)
21 cities in the world (all landmarks)	25.5	26.4	+0.9
21 cities in the world (top 25 landmarks)	29.3	33.4	+4.1
21 cities in the world (top 10 landmarks)	29.0	39.3	+10.3

ranges and the transition probability matrices from the current location  $L_i$  to the next location  $L_j$  are generated independently from each other season and time.

Table 10.1 shows the accuracy comparison of the intra-city landmark recommendation. Although the proposed model is simple, it is demonstrated that the performance improvement is much larger than the other approaches [8, 22]. Note that the datasets are different from each other because there is no open dataset. Cheng's work [8], in particular, requires that faces can be detected in the photos, therefore they require different dataset from [22] and ours. It is also observed that the performance becomes better when the number of landmarks to consider is smaller. This observation does not necessarily mean that travel navigation becomes easier with less number of landmarks to consider. In fact, if we look at the results by the naive Markov approach, the navigation accuracies are almost all the same regardless of the number of landmarks. This indicates that the difficulty of travel navigation is almost independent of the number of landmarks but popular landmarks tend to have stronger seasonal/temporal dependence in their popularity.

Two interesting examples from specific cities (New York and Taipei) are shown in Fig. 10.4 and Table 10.2. Here, the results considering the top 25 landmarks are presented. It is shown that the recommendation accuracy is improved in both cities. The accuracy becomes much better in New York (from 12.1 to 22.8 %) because some landmarks are popular at a certain season or time of the day. For example, the Rockefeller center gets more popular and the Statue of liberty gets less popular, respectively, in winter though these landmarks are always popular in all seasons. On the other hand, it can be observed that landmarks in Taipei have less such season or time specific popularity, indicating that the popularities of the landmarks in Taipei are independent



**Fig. 10.4** Recommendation accuracy for specific cities: **a** New York **b** Taipei. The top 25 landmarks were considered

**Table 10.2** Accuracy of intra-city recommendation in specific cities

	Naive Markov (%)	Our results (%)	Improvement from Markov (%)
New York	12.1	22.8	+10.7
Taipei	35.4	37.8	+2.4



**Fig. 10.5** “Where to go” navigation for photo-shooting

of time and season. In fact, popular landmarks in Taipei such as Taipei 101, National Palace Museum, etc., are indoor attractions and less sensitive to season and time.

#### **10.5.4 Photo-Shooting Navigation**

Figure 10.5 demonstrates the “where to go” navigation examples. In Fig. 10.5, a landscape photo taken at the Tokyo tower is used as a query and the recommended landmarks in Paris where similar photos can be taken are displayed on the map along with the photo examples. In Fig. 10.5, two different cherry blossoms are used as queries and the navigation system properly suggests where to visit in order to take similar photos.

In Fig. 10.6, results of user-to-user-similarity-based navigation are shown. Here, the user inputs three photos he/she has taken in Paris and the system retrieves the users who have taken similar photos. In Fig. 10.6, the query images and a retrieved user and his photos are shown. Then, the other photos that the retrieved user has taken in Paris are displayed on the map. In this manner, the system can navigate where to visit depending on the user’s photo-shooting styles.

Examples of the “how to shoot” navigation are demonstrated in Fig. 10.7. The photos of the Eiffel tower in Paris and the Rockefeller center in New York with higher attractiveness score are shown in Fig. 10.7a and b, respectively. By showing such photo examples, the users can be advised how to take good photos by themselves.



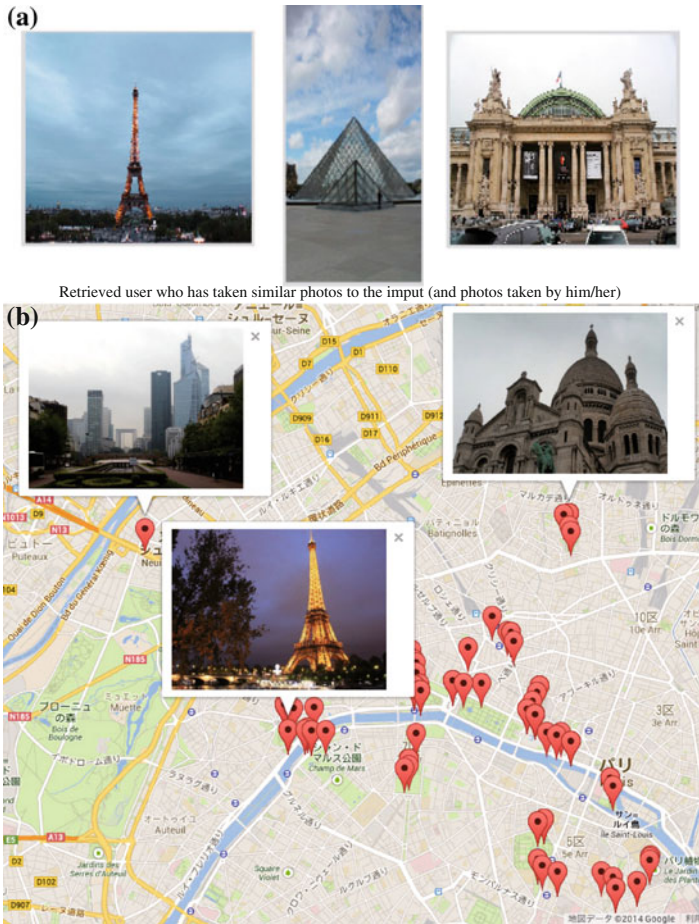


Fig. 10.6 “Where to go” navigation based on similar user retrieval based on photo preference: a query and retrieved photos and b suggested landmarks and photos

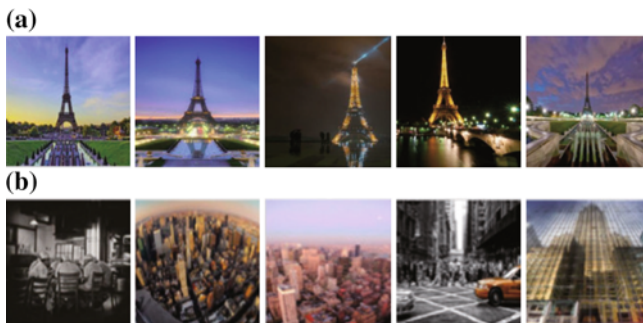


Fig. 10.7 “How to shoot” navigation: a Eiffel tower in Paris. b Rockefeller center in New York

## 10.6 Conclusions

This chapter presented a personalized travel and photo-shooting navigation algorithms. Our travel navigation model featuring seasonal and temporal information can improve the recommendation accuracy better than previous approaches. It is also possible to combine our proposed algorithm with the previous travel navigation approaches. Experiments using more than 6.2 million geo-tag data demonstrated the validity of our proposed algorithm. In addition, photo-shooting navigation examples using large-scale geotagged photos have also been presented.

## References

1. A. Agarwala, Efficient gradient-domain compositing using quadtrees. *ACM Trans. Graph.* **26**(3), 94 (2007)
2. Y. Arase, X. Xie, T. Hara, S. Nishio, Mining people's trips from large scale geo-tagged photos, in *ACMMM*, pp. 133–142 (2011)
3. C. Barnes, E. Shechtman, A. Finkelstein, D.B. Goldman, PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Trans. Graph. (Proc SIGGRAPH)* **28**(3), (2009)
4. M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, Image inpainting, in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00*, pp. 417–424 (2000)
5. S. Bhattacharya, R. Sukthankar, M. Shah, A framework for photo-quality assessment and enhancement based on visual aesthetics, in *ACMMM*, pp. 271–280 (2010)
6. M. Birsak, P. Musialski, P. Wonka, M. Wimmer, Automatic generation of tourist brochures. *Computer Graphics Forum (Proceedings of EUROGRAPHICS 2014)* **33** (2014)
7. W.C. Chen, A. Batesttini, N. Gelfand, V. Setlur, Visual summaries of popular landmarks from community photo collections, in *ACMMM*, pp. 789–792 (2009)
8. A.J. Cheng, Y.Y. Chen, Y.T. Huang, W.H. Hsu, H.Y.M. Liao, Personalized travel recommendation by mining people attributes from community-contributed photos, in *ACMMM*, pp. 83–92 (2011)
9. J. Choi, G. Friedland, V. Ekambaram, K. Ramchandran, Multimodal location estimation of consumer media: Dealing with sparse training data, in *ICME*, pp. 43–48 (2012)
10. M.D. Choudhury, M. Feldman, S. Amer-Yahia, N. Golbandi, R. Lempel, C. Yu, Constructing travel itineraries from tagged geo-temporal breadcrumbs, in *WWW*, pp. 1083–1084 (2010)
11. M. Clements, P. Serdyukov, A.P. de Vries, M.J. Reinders, Using flickr geotags to predict user travel behaviour, in *SIGIR*, pp. 851–852 (2010)
12. D. Crandall, L. Backstrom, D. Huttenlocher, J. Kleinberg, Mapping the world fs photos, in *WWW*, pp. 761–770 (2009)
13. Q. Fang, J. Sang, C. Xu, Giant: Geo-informative attributes for location recognition and exploration, in *ACMMM*, pp. 13–22 (2013)
14. F. Girardin, F. Calabrese, F. Fiore, C. Ratti, J. Blat, Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Comput.* **7**(4), 36–43 (2008)
15. J. Hays, A. Efros, Im2gps: estimating geographic information from a single image, in *CVPR*, pp. 1–8 (2008)
16. J. Hays, A.A. Efros, Scene completion using millions of photographs. *ACM Trans. Graph.* **26**(3), 87–94 (2007)
17. N. Ishihara, Y. Itoh, K. Takashima, F. Kishino, Attractiveness-based image scoring using similar images on websites, in *IPSIJ Interaction (in Japanese)* (2011)

18. A. Jaffe, M. Naaman, T. Tassa, M. Davis, Generating summaries and visualization for large collections of geo-referenced photographs, in *ACM MIR*, pp. 89–98 (2006)
19. K. Jiang, P. Wang, N. Yu, Contextrank: Personalized tourism recommendation by exploiting context information of geotaggedweb photos, in *ICIG*, pp. 931–937 (2011)
20. L. Kennedy, M. Naaman, Generating diverse and representative image search results for landmarks, in *WWW*, pp. 297–306 (2008)
21. L. Kennedy, M. Naaman, S. Ahern, R. Nair, T. Rattenbury, How flickr helps us make sense of the world: context and content in community-contributed media collections, in *ACMMM*, pp. 631–640 (2007)
22. T. Kurashima, T. Iwata, G. Irie, K. Fujimura, Travel route recommendation using geotags in photo sharing sites, in *CIKM*, pp. 579–588 (2010)
23. Li C, Loui AC, Chen T (2010) Towards aesthetics: A photo quality assessment and photo selection system, in *ACMMM*, pp. 827–830
24. J. Li, Z. Qian, Y.Y. Tang, L. Yang, T. Mei, Gps estimation for places of interest from social users' uploaded photos. *IEEE Trans. Multimedia* **15**(8), 2058–2071 (2013)
25. Z. Luming, S. Mingli, Y. Yi, Z. Qi, Z. Chen, N. Sebe, Weakly supervised photo cropping. *IEEE Trans. Multimedia* **16**(1), 94–107 (2014)
26. Mamei M, Rosi A, Zambonelli F (2010) Automatic analysis of geotagged photos for intelligent tourist services, in *Sixth International Conference on Intelligent Environments*, pp. 146–151 (2010)
27. Nishiyama M, Okabe T, Sato Y, Sato I (2009) Sensation-based photo cropping, in *ACMMM*, pp. 669–672
28. Popescu A, Grefenstette G (2011) Mining social media to create personalized recommendations for tourist visits, in *COM.Geo*
29. A. Popescu, G. Grefenstette, P.A. Moëllic, Mining tourist information from user-supplied collections, in *CKIM*, pp. 1713–1716 (2009)
30. Quack T, Leibe B, Gool LV (2008) World-scale mining of objects and events from community photo collections, in *CIVR*, pp. 47–56
31. T. Rattenbury, M. Naaman, Methods for extracting place semantics from flickr tags. *ACM Trans. Web* **3**(1), 1:1–1:30 (2009)
32. H. Shu, C. Chen, T. Chen, Landmark recognition: A unary approach, in *ICIP*, pp. 2645–2648 (2010)
33. J. Sun, L. Yuan, J. Jia, H. Shum, Image completion with structure propagation. *ACM Trans. Graph.* **24**(3), 861–868 (2005)
34. J. Wu, J. Rehg, Centrist: A visual descriptor for scene categorization. *IEEE TPAMI* **33**(8), 1489–1501 (2011)
35. Yamasaki T, Gallagher A, Chen T (2013) Personalized intra- and inter-city travel recommendation using large-scale geo-tagged photos, in *Geomm*
36. Yeh CH, Ho YC, Barsky BA, Ouhyoung M (2010) Personalized photograph ranking and selection system, in *ACMMM*, pp. 211–220 (2010)
37. Yin H, Lu X, Wang C, Yu N, Zhang L (2010) Photo2trip: an interactive trip planning system based on geo-tagged photos, in *ACMMM*, pp. 1579–1582
38. Yunpeng L, Crandall D, Huttenlocher D (2009) Landmark classification in large-scale image collections, in *ICCV*, pp. 1957–1964
39. Y. Zheng, S. Yan, Z. Zha, Y. Li, x Zhou, T. Chua, R. Jain, Gpsview: A scenic driving route planner. *ACM Trans. Multimedia Comput. Commun. Appl.* **9**(1), 3:1–3:18 (2013)