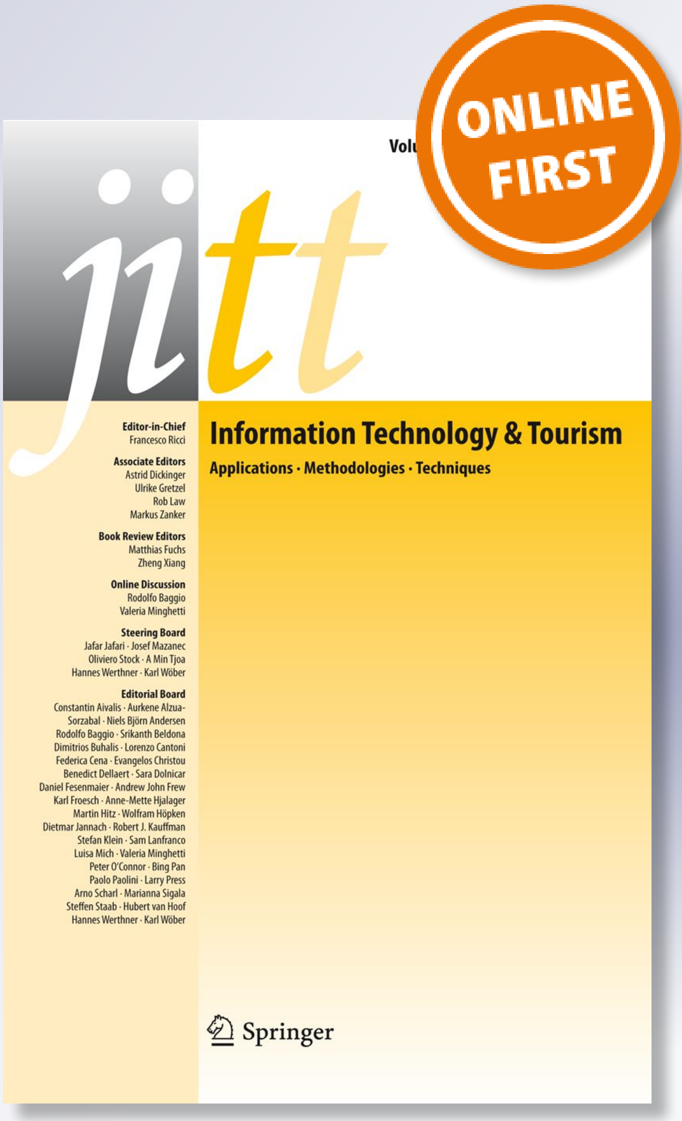
Understanding tourists' photo sharing and visit pattern at non-first tier attractions via geotagged photos

Rosanna Leung, Huy Quan Vu & Jia Rong

Information Technology & Tourism

ISSN 1098-3058

Inf Technol Tourism
DOI 10.1007/s40558-017-0078-3



Springer

Inf Technol Tourism
DOI 10.1007/s40558-017-0078-3

ORIGINAL RESEARCH

Understanding tourists' photo sharing and visit pattern at non-first tier attractions via geotagged photos

Rosanna Leung[1] · Huy Quan Vu[2] · Jia Rong[2]

Received: 23 June 2016 / Revised: 12 December 2016 / Accepted: 17 February 2017

Abstract Social media plays an important role in tourism industry, especially for individual travel planning and tourism entities preparing business plans. Only a limited number of first-tier attractions were reported in tourism bureau's travel statistics documents, which cannot satisfy the needs of non-first tier attraction managers preparing their marketing strategies. With the rich tourists reviews and photos publicly available on social network platform, researchers and attraction manager could analyzing these geotagged photos to find out the potentials of the attractions including tourists interests and their travel pattern. In this study, we report our work on extracting and processing of geotagged photos uploaded by inbound tourists on Flickr.com to study tourists' photo sharing and visiting pattern during their visits at Hong Kong temples. Four popular temples were identified automatically using P-DBSCAN density clustering from geotagged tourists photos. The travel pattern analysis had shown that tourists from different country of residence have different temple choice. Particularly, a closer look at the repeated

This paper is an extended and updated version of a conference paper titled 'Tourists Visit and Photo Sharing Behavior Analysis: A Case Study of Hong Kong Temples' previously published in the proceedings of Information and Communication Technologies in Tourism 2016 Conference (ENTER 2016) held in Bilbao, Spain, February 2–5, 2016.

✉ Rosanna Leung
rosannaleung@isu.edu.tw

Huy Quan Vu
huyquan.vu@vu.edu.au

Jia Rong
jia.rong@vu.edu.au

[1] Department of International Tourism and Hospitality, I-Shou University, Kaohsiung, Taiwan

[2] Centre of Applied Informatics, College of Engineering and Science, Victoria University, Melbourne, Australia

Published online: 06 March 2017

tourists in the past five years, and special focus on photo uploading habits are discussed in our findings.

Keywords Temple · Hong Kong · Geotagging · P-DBSCAN clustering · Photo sharing behavior · Visit pattern

1 Introduction

Increasing number of tourists sharing their travel diary on online media-sharing platforms such as Facebook, Youtube, Instagram and Flickr, provided massive amount of photos and videos on the Internet, and have become an important data source for both academic and industry researches. Scholars and attraction managers could make use of this large number of publicly available information such as user profiles, textual and visual context upload by the tourists to examine the tourist visit pattern and their travel preferences. Their selections of the uploaded photos reflect their personal opinions of the past travels, and perceptions about the destination which can directly or indirectly affect the peers' impressions and decision-making of the next trip (González-Rodríguez et al. 2016). Most of the latest digital cameras and smart phones have built-in GPS features so that when people taking photos with such devices, the photos would be "geotagged" with the geographical locations generated. *GeoTagging* (here after geotag) is a process of attaching geographical identification metadata into multimedia files such as images and videos and identify tourists behavior and movement (Zheng et al. 2012). Social media sites enable anyone to retrieve and extract photos based on the geotag information so even the tourist does not type in any description for the photos, the users' profile and the photos themselves can provide certain details for visit pattern analysis. Attraction managers therefore can analyze such information to get better understanding of tourists' preferences and behavioral differences. This study aims to identify the popular attractions spots in Hong Kong, and analyze tourists' visit pattern using geotagged photo shared on social media sites together with the associated Metadata.

With only 1104 km^2, receiving 60 million incoming tourists (HKTB 2016) and more than 27 million international visitors, Hong Kong has ranked the top popular city around the world in 2014 (Euromonitor International 2015). Hong Kong Tourism Board (HKTB) has been promoting Hong Kong as a shopping and dining paradise. Their annual reports illustrated detail tourist expenditure, and tourist statistics of a range of tourist attractions. As these statistics only focused on primary and selected secondary attractions; any places beyond the study scope are excluded due to time and budget constraints for conducting comprehensive survey to gather information from every tourist. As a result, those tourists' activities without monetary value or secondary and tertiary tourist attractions (hereafter refer as non-first tier attractions) were unable to obtain any statistics as from tourism board. Moreover, tourists from different regions with different cultural background behave differently (Leung et al. 2012). Attraction managers have difficulties to identify their positioning and categorize data for geographical analysis. They can only rely on their own observations from limited number of samples, or information prepared by others to design their marketing and business strategies (Lew and McKercher 2006). According to HKTB (2014), visiting temples was one of the cultural tourists' activities in Hong Kong with around 7% of the tourists visited temples in Hong Kong (HKTB 2014). However, no prior research or statistical report indicates which temples were in the top visited list so this study would attempt to use tourist photos to assist non-first tier attractions to identify the top attraction in their sectors and collect tourists' statistics from online media-sharing website.

The rest of paper is structured as follows: Sect. 2 reviews existing literature on analyzing tourist travel pattern using GPS and geotag information. The proposed methodology for collecting and processing geotag data attached with the tourist uploaded photos is introduced in Sect. 3. Section 4 presents a case study of analyzing tourists' visits pattern and photo sharing behaviors in Hong Kong temples. Section 5 concludes the reported work and emphasizes on the current limitations.

2 Literature review

Destination photos could shape and reshape tourist's perception about destination and influence their process of decision making. Since Geographic Information System (GIS) and Global Positioning System (GPS) introduced, it enable researchers to explore tourists' movement patterns (Lau and McKercher 2006; Leung et al. 2012). A common way to conduct research on tourist movement required tourists to carry a GPS enable device to record their travel movement though out the trip. These kind of research could precisely capture the actual tourists' path with all the locations they have visited, but with many limitations including: (1) small samples size due to limited number of devices available; (2) time consuming to record the entire trip; and (3) low participation rate due to the inconvenience caused from carrying the device or privacy concern from capturing the travel route. With the build-in feature of Global Positioning System (GPS) in smart phones and mobile photo capturing devices, the geographical information can now automatically stored in a photo's geotag for location recording. In order to make use of those geotagged photos available on Internet, many automatic computing algorithms were developed to collect, store and organize those photos for further analysis such as location clustering (Yin et al. 2011), and visual tag maps (Yanai et al. 2009). Context on the travel photos can represent photographer's own mental image and reflect the inner feeling of the destination (Crawshaw and Urry 1997; Pan et al. 2014) therefore analyzing tourists' photos can have a clear picture about the destination in through the eyes of the tourists. Stepchenkova and Morrison (2006) manually categorized the tourists' photos from Flickr and blogs; examine the photo image by content analysis, and attempt to identify the cultural difference among tourists. Stepchenkova and Zhan (2013) examined the geographic characteristics of Peru from photos uploaded on Flickr and compared the images differences created by DMO and tourists. Pan et al. (2014) extracted the captions of travel photos and comments for a pre-determined destination to evaluate destinations' quality, relationships among travel motivations, resolutions of images taken at that destination, and other affective qualities. However, all these studies were conduct manually from data collection to data analysis which limited the numbers of photos that can be analyze.

Crowdsourced data from photo-sharing platforms have the potential to serving as a proxy of space attractiveness (Kachkaev and Wood 2013). By using large number of collected geotagged photos, attraction managers able to analyze and understand the tourists' visit behaviors and travel patterns for future marketing planning. Zheng et al. (2012) extracted photos in four modern cities (London, Paris, San Francisco, and New York City) and analyzing tourist movement patterns in relation to region of attractions. By capturing more than 1 million photos with text description related to the name of the cities in Austria from Flickr, Önder et al. (2014) traced the travel pattern of tourists travelling in Austria. García-Palomares et al. (2015) make use of 43,000 tourists' photos from *Panoramio.com* and identified tourist hot-spot in four cities in Europe. Many recent studies focused on analyzing textual information attached to the uploaded photos, such as hash tags (Giannoulakis and Tsapatsoulis 2015), photo captions (Pan et al. 2014), and review comments (Hall et al. 2015). However, if the photo does not associate to any textual context, it is not easy to adopt these methods and obtain expected outcome. Kisilevich et al. (2010) ranked of sightseeing places in a city using geotagged photos. Vu et al. (2015) attempted to use geotagged data to analyze the tourists' travel preferences in Hong Kong. Their study identified the most popular attraction sites in Hong Kong, and the differences in travel patterns and the photo sharing behavior between Asian and Western tourists. However, none of these studies focused on non-first tier tourist attraction. Industry practitioners at second or third tier attractions have to struggle with data they need when prepare for strategic plans. This study determined to fill the gaps by exploring the tourists' travel patterns at the non-first tier tourist attractions such as temple.

3 Methodology

This section presents our approach to explore the popular temples in Hong Kong. An issue with the analysis is that a tourism destination may have many temples. Some temples may not have clear boundaries with outside areas, while, visitors may take temple photos from either inside or outside of the temple. It is necessary to identify the most popular ones and their geographical boundary to focus the data extraction and analysis. We propose to perform the analysis in three steps: (1) keyword-based geotagged photo searching and initialized analysis; (2) popular temple identification; and (3) geotagged photo extraction for identified popular temples and visit pattern analysis.

3.1 Keyword-based geotagged photo searching and initialized analysis

Geotagged photos are available on the web applications of Flickr for public view, but they were not directly downloadable. They must be accessed via Flickr's Application Programming Interface (API), whose documentation is given in (Flickr

2015). Among the wide ranges of functions provided by Flickr's API, *PhotosSearch* function allows users to query Flickr's servers and retrieve information based on certain search criteria. The location of each geotagged photo p is referenced by a value pair $<x_p, y_p>$ for longitude and latitude coordinates. The region defined to extract geotagged photos can be specified by a bounding box, whose coordinates are defined by $x_{min}, y_{min}, x_{max}$ and y_{max} for the minimum longitude, minimum latitude, maximum longitude, and maximum latitude, respectively. The bounding box coordinates for a tourism destination can be determined using Google Map (www.google.com.au/maps).

Besides, a number of specific properties can be set to obtain photos and corresponding Metadata with various purposes, such as searching photos with special keyword(s); setting up certain time period of photo taking. The keyword(s) can be defined by a string of text with one or more words, describing entity of interest. The *taken time* interface can be specified by t_{min} for the earliest time and t_{max} for the latest time. Only photos taken between these periods are considered. The returned result contains Metadata information carried by the photos including *PhotoID*, *GPS location*, *TakenDate*, and *Tags*.

3.2 Popular temples identification

Geotagged photo data set were collected from Flickr with specific keywords such as "temple", the next step is to identify the most popular temples and define the scope for data analysis. It is possible that a temple might be visited by few tourists but took many photos, while another temple is visited by many tourists but each took few photos. The popularity of a temple should be determined based on the number of tourists rather than the number of photos taken. We adopted a density clustering technique, named P-DBSCAN, to assist identifying the popular temple (Kisilevich et al. 2010). The advantage of P-DBSCAN is the ability to account for both user numbers and photo numbers in its computation, whose details are given below.

Suppose P is a collection of geotagged photo data relevant to temple. Distance between two photo p_i and p_j is defined as $Dis(p_i, p_j)$. Let r be a neighborhood radius. The neighborhood photos $N_r(p_i)$ of a photo p_i is defined by:

$$N_r(p_i) = (p_j \in P', Owner(p_j) \neq Owner(p_i) | Dis(p_i, p_j) \leq r), \qquad (1)$$

where $Owner(.)$ is an ownership function to specify the owner of photo p_i. Let $NeighborOwner(p_i)$ be the owner number of the neighbor photo $N_r(p_i)$, and δ be a owner number threshold. Photo p_i is called a core photo if $NeighborOwner(p_i) \geq \delta$. At the beginning of the clustering process, all photos are marked as unprocessed. For each photo p_i, if it is a core photos, it is assigned to a cluster c and its neighbors are assigned to a queue to be processed next; otherwise, p_i is discarded. Each of the neighboring photos is then processed and assigned to the current cluster c until the queue is empty. The process is iterated for the rest of the photo in P, and result a set of clusters C. The values r and δ are determined based on the scale of specific applications. If the region to be identified is at the macro level such as a country or a city, large values can be assigned to r and δ. If the region is at a micro level such as a temple, r and δ take small values. The geographical coordinates of the clusters are then examined to determine the location and name of the temples. Note that a criteria in the clustering process is the owner number, the cluster is identified based on the actual number of visitors rather than the number of photos. This ensures the identified clusters are the popular location for tourist visits.

3.3 Geotagged photo extractions for identified popular temples and visitor pattern

Although, the photos collected in the previous steps help identify the popular temples, there are cases that some photos taken inside a temple were not tagged with any relevant keyword, and some photos with the "temple" tag were not actually taken inside a temple. A second round of data collection is conducted in order to extract all photos taken at the locations identified as popular temples. Based on the clustering result, the geographical areas of the selected temples are determined. A set of new bounding boxes are then designed to cover entirely the interested areas. *PhotosSearch* function is applied again, but no keyword is required so that all photos taken in the certain areas are collected.

In this second round of data collection, all the photos are downloaded with the complete metadata, including *GPS location*, *TakenDate*, *OwnerID* and owner's *Location of Resident*, especially having an extra feature *UploadDate*. The *uploadDate* is used to study the customer sharing behaviors that could indicate the timeliness of the uploaded content. If a photo was uploaded a long time after it was taken, then it may not be able to show the current situation at the destination.

With the collected information, a series of analysis is carried out using descriptive statics. Namely, geographical distribution of temple visitors is determined by counting the number of users from each location. Visiting patterns are computed by count the number of users according to year, month and day based on the *TakenDate*. For each tourist, we can determine the number of visited temples, or the number of visits to each temple based on the *OwnerID*. A tourist is known to visit two temples if photos with the same *OwnerID* appear in the geographical area of two different temples. The actual content of the photos is examined to determine the context and interests of visitors at different temples.

4 Experiment implementation and finding analysis

The proposed methodology was adopted to analyze temple visitor pattern in Hong Kong. In this section, the process of experiments and the obtained corresponding results are presented and discussed to provide suggestions and advices for management in tourism industry for marketing purpose.

4.1 Popular temple identification

To identify the popular temples for further analysis, we set the bounding box with parameter values shown in Table 1 that covered the entire Hong Kong geographical area, as suggested in recent work (Vu et al. 2015). The search was limited to recent five and half years from January 1, 2010 till June 30, 2015. Keywords inputted into the search function were "temple" and "buddha" due to most of the temples in Hong Kong were built up for Buddhism. If the photo tag field contained one of the provided keywords, then certain photo was included in the returned results; otherwise, it was discarded.

The search returned 3767 photos related to temples from 783 users over the entire Hong Kong area. The locations of the collected photos are shown as yellow dots on the satellite image as Fig. 1. P-DBSCAN was then applied to the collected dataset for clustering. In our case, the regions of interest were the temples at micro level, thus, r can take small values of 0.002 as recommended in Vu et al. (2015), which is equivalent to approximately 150 m. The minimum owner δ was set to 5% of the total number owners in data collection. The clustering process returned four clusters as shown in Fig. 1b. After examining the locations, the clusters returned were marked using the names of the corresponding temples. These temples were *Tian Tan Buddha*, *Wong Tai Sin, Tin Hau* and *Man Mo*.

4.2 Geotagged photos collection for the identified temples

From Fig. 2a, b, we found that the photos taken at the Tian Tan Buddha and Wong Tai Sin clusters were mainly located within or close to the regions belonging to those temples. In this case, we can easily define bounding boxes (as shown in red rectangular), to cover the spatial extends of these temples to extract the data for further analysis. For *Tin Hau* cluster (Fig. 2c), many photos were indeed not taken at Tin Hau Temple, but along the "Temple Street". This is due to the specification of "temple" keyword during the photo search process crash with the street name. Many photos at "Tin Hau Temple Street" were included but indicated as irrelevant to our interest of temple visitors. However, "Tin Hau Temple Street" was named after Tin Hau temple. Therefore, Tin Hau temple was still included by setting a bounding box only covered the area belongs to the temple. For the cluster shown in Fig. 2d, the photos were centered at Man Mo temple, but spread widely to the surrounding areas. The reason is because Man Mo temple stays in the center of Hong Kong metropolitan area, and was surrounded by large number of skyscrapers. The accuracy of GPS location calculation might be affected by the signal reflection

Table 1 Photo search parameters

Parameter	Value	Description
x_{min}	113.887603	Minimum longitude of the bounding box
y_{min}	22.215377	Minimum latitude of the bounding box
x_{max}	114.360015	Maximum longitude of the bounding box
y_{max}	22.51446	Maximum latitude of the bounding box
t_{min}	1/1/2010	Earliest photo taking date
t_{max}	30/6/2015	Latest photo taking date

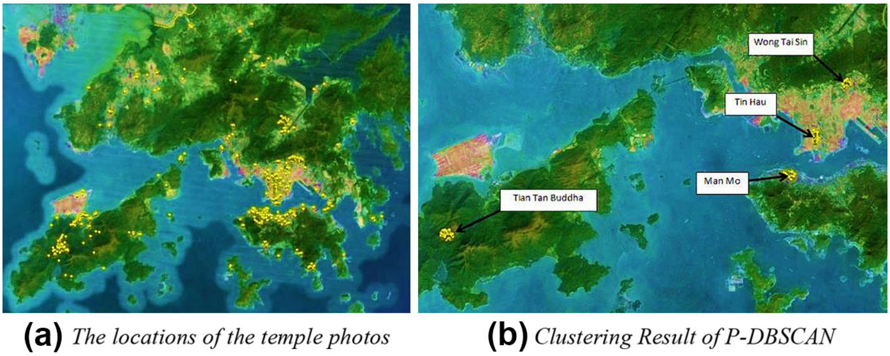


Fig. 1 Locations of temple photos

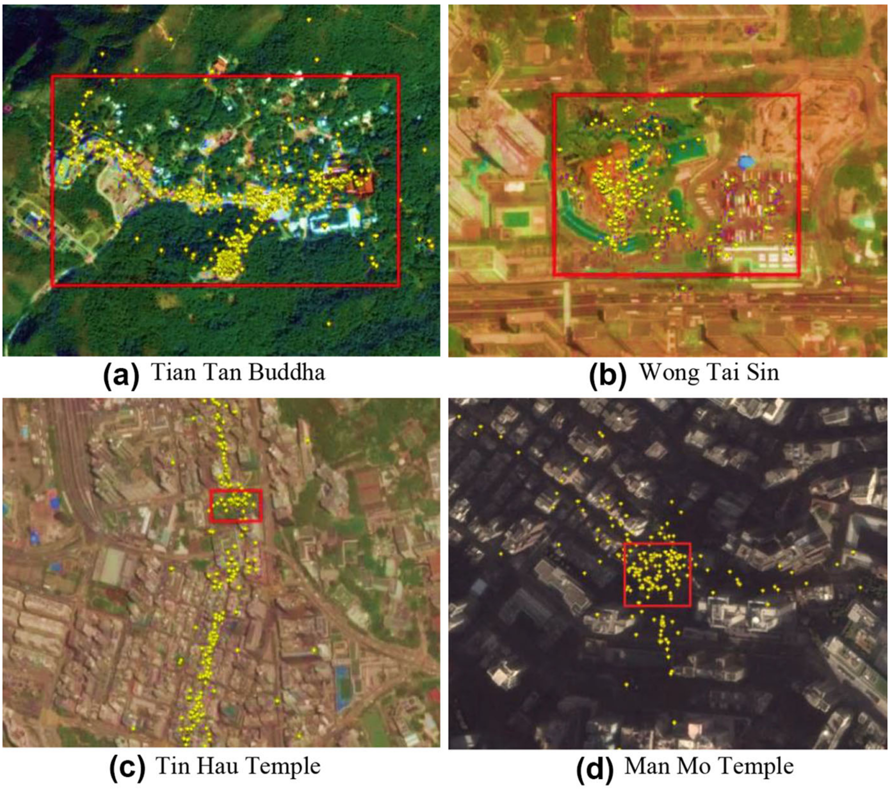


Fig. 2 Locations of photos for temple clusters

from these buildings. In order to extract the photos that were taken at the Man Mo temple, this study set the bounding box to cover the area around Man Mo temple only, which was around the center of the cluster.

A second round data collection was conducted for the four popular temples (Fig. 2). To avoid photos taken at the temple that were not tagged with any relevant keyword, a new data set with all the geotagged photos taken inside the bounding boxes for *Tian Tan Buddha*, *Wong Tai Sin, Tin Hau* and *Man Mo* temples were extracted. Totally 6956 photos from 780 visits were returned. Table 2 shows the photos taken at each temple and its corresponding number of visits. Tian Tan Buddha was the most popular one with around 5000 photos taken from there by more than 500 tourists. Wong Tai Sin was the second popular temple with 1372 photos. However, the average number of photos uploaded per visit was the highest among all four temples.

4.3 Geographical differences on tourists' temple visits

Flickr enable users to fill in their residential country in their profile. In this study, out from the 734 visitors, 47% of them (318 visitors) indicated their country of residence in the profile. One-third of the temple visitors were from Europe, 28% from Asia, 19% from North America, and 11% from both Australia/New Zealand and South America. After 2013, the numbers of users that uploaded temple photos were dropped. There was no solid evidence to proof the reason of the dropping. It could relate to Flickr's popularity, the dropping of visitors' interest on temples, or the dropping of tourists' photo sharing behavior. Table 3 indicates the frequency distribution of Flickr users that had uploaded Hong Kong temple photos by region. By calculating the average number of photos upload per region, North America ranked the first with 14 photos per tourists. Asia and Australia/New Zealand ranked second with nine photos per tourists, and Europe ranked third with eight photos per person. As there were 17 visitors with multiple temple visits, the total number of visits by region was 335.

As shown in Table 4, around 51 tourists (17%) were from USA; UK ranked second (12.54%) and Mainland China tourists ranked third (8.75%) with 41 and 26 visitors respectively. Other than overseas tourist, 9% of the temple visitors were local residents (Table 3). Out from all 36 countries, only seven tourists from five countries repeat visiting the temples in Hong Kong in the past 5 years. Two each were from USA and China, and one each from Japan, Korea and Hong Kong. The number of photos uploaded (shown in the blankets after the number of the visitors) indicated Indian on averaged uploaded 23 photos per tourists; Russian ranked

Table 2 Statistics of Tourists and Photos Uploaded for the Popular Temples

Popularity	Temples	Photos	No. of visits	Average photos uploaded per visit
1	Tian Tan Buddha	4965	541	9.18
2	Wong Tai Sin	1372	126	10.88
3	Man Mo Temple	363	60	6.05
4	Tin Hau Temple	256	53	4.83
	Total	6956	780	8.92

Table 3 Flickr User with Temple Photos Uploaded by Region

Region	Tian Tan	Wong Tai Sin	Man Mo	Tin Hau	Total no. of visits	Unique visitor[a] (avg photo per visitor)
Europe	86 (732)	11 (90)	12 (49)	6 (14)	115 (885)	109 (8)
Asia	57 (568)	14 (100)	12 (64)	9 (34)	92 (766)	86 (9)
North America	45 (613)	12 (192)	4 (27)	2 (15)	63 (847)	59 (14)
Australia/New Zealand	22 (175)	1 (3)	1 (5)	1 (49)	25 (232)	25 (9)
South America	9 (37)	1 (7)	1 (6)	–	11 (50)	11 (5)
Hong Kong	15 (119)	6 (58)	3 (10)	5 (19)	29 (206)	28 (7)
Region Total	234 (2244)	45 (450)	33 (161)	23 (131)	335 (2986)	318 (9)
Not stated	307 (2721)	81 (922)	27 (202)	30 (125)	445 (3970)	416 (10)
Total	541 (4965)	126 (1372)	60 (363)	53 (256)	780 (6956)	734 (9)

Numbers in bracket indicated the number of photos uploaded

[a] One visitor could have multiple temple visits

second with 21 photos per person, and USA and Korean ranked third with 15 photos per tourist.

4.4 Repeated visiting pattern analysis and multiple temple visits

The total number of visits collected in this study was 780 (Table 2) but the actual number of Flickr visitors was 734 as one person can visit multiple temples in one trip (Table 5). By looking at the visitor travel patterns in the identified four temples, none of the visitors have visited all four of them. Only five tourists visited three temples, and 36 of them visited two. The remaining 94% tourists only visited one temple, and majority of them went to see the Tian Tan Buddha. Tian Tan received 541 visits and ranked as the most popular attractions among temple tourists, followed by Wong Tai Sin with 126 visitors. Tin Hau and Man Mo temples had 60 and 53 visits respectively. Table 5 presents the statistics of tourists' temple visits.

In order to have a clearer picture of the tourists' behavior, we extracted a set of users who visited multiple temples and/or visited temples in multiple years and summarized their visits in Table 6. Out from 780 visitors, 41 of them (5%) matched these criteria and 17 of them indicated the country of residence in their profile. Out of these 41 visitors, only nine of them visited Hong Kong temples in different years. One of them was Hong Kong resident, three of them were from Asia, and the remaining did not indicate their country of residence. Tourist1 have a total of five temple visits in 2012 and 2014 to three temples. In the first visit, he/she has visited three temples; and two temples in the second visit. From the photo count, we can presume this tourist may not interested in Tin Hau temple as there was only one photo upload and he/she did not revisit it in the second trip. For Tourist2, he/she consecutively visit Hong Kong in 3 years. The repeat visits only visited Tian Tan so

Table 4 Geographical distribution of base on flickr user profile by year

Country	2010	2011	2012	2013	2014	2015[b]	Visits (total no. of photos)	Visitors (avg photo per visitor)
USA	14 (263)	6 (85)	10 (50)	11 (210)	8 (150)	2 (5)	51 (763)	51 (15)[a]
UK	7 (69)	12 (73)	11 (42)	3 (14)	5 (145)	3 (13)	41 (356)	41 (9)
China	4 (116)	7 (23)	9 (79)	4 (18)	3 (6)	2 (2)	29 (244)	26 (9)[a]
Australia	4 (19)	5 (94)	3 (5)	5 (27)	1 (45)	2 (18)	20 (208)	20 (10)
Singapore	2 (21)	5 (41)	1 (4)	6 (92)	3 (38)	–	17 (196)	17 (12)
Germany	3 (61)	4 (50)	–	3 (7)	2 (12)	2 (6)	14 (136)	14 (10)
Taiwan	1 (12)	2 (43)	5 (14)	3 (7)	1 (15)	1 (1)	13 (92)	13 (7)
Japan	4 (64)	2 (18)	1 (2)	2 (5)	–	–	9 (89)	8 (11)[a]
Spain	3 (17)	1 (14)	1 (3)	1 (5)	2 (3)	1 (7)	9 (49)	9 (5)
Canada	1 (28)	1 (2)	2 (3)	3 (13)	1 (38)	–	8 (84)	8 (11)
Russia	–	2 (11)	5 (53)	1 (102)	–	–	8 (166)	8 (21)
Philippines	3 (7)	1 (5)	–	1 (18)	2 (17)	–	7 (47)	7 (7)
Italy	1 (1)	3 (7)	1 (1)	1 (1)	–	–	6 (10)	6 (2)
Netherlands	–	1 (3)	3 (42)	1 (2)	1 (11)	–	6 (58)	6 (10)
Switzerland	–	1 (3)	–	2 (15)	3 (12)	–	6 (30)	6 (5)
Brazil	–	1 (3)	–	2 (8)	1 (7)	1 (6)	5 (24)	5 (5)
France	3 (8)	1 (2)	–	1 (12)	–	–	5 (22)	5 (4)
Malaysia	2 (7)	–	–	1 (1)	2 (2)	–	5 (10)	5 (2)
New Zealand	1 (5)	1 (4)	2 (10)	1 (5)	–	–	5 (24)	5 (5)
Thailand	–	–	–	1 (6)	2 (3)	1 (3)	4 (12)	4 (3)
Argentina	–	2 (9)	–	1 (1)	–	–	3 (10)	3 (3)
Belgium	–	1 (3)	–	1 (6)	1 (6)	–	3 (15)	3 (5)
Sweden	2 (7)	1 (2)	–	–	–	–	3 (9)	3 (3)
Finland	–	–	–	2 (9)	–	–	2 (9)	2 (5)
India	–	1 (41)	–	–	1 (4)	–	2 (45)	2 (23)
Korea	–	1 (2)	–	–	–	1 (13)	2 (15)	1 (15)[a]
Mexico	–	–	1 (6)	1 (2)	–	–	2 (8)	2 (4)
UAE	–	1 (1)	–	–	–	1 (2)	2 (3)	2 (2)
Others	1 (1)	1 (13)	1 (1)	3 (21)	1 (2)	1 (8)	8 (46)	8 (6)
Hong Kong	5 (96)	9 (35)	3 (18)	9 (42)	3 (8)	2 (7)	31 (206)	28 (7)[a]
Region Total	62 (802)	73 (587)	59 (333)	71 (649)	43 (524)	20 (91)	328 (2986)	318 (9)
Not Stated	66 (428)	64 (490)	73 (553)	102(1513)	87 (762)	39 (224)	431 (3970)	416 (10)[a]
Total	128 (1230)	137 (1077)	132 (886)	173 (2162)	130 (1286)	59 (315)	759 (6956)	734 (9)

Numbers in bracket indicated the number of photos uploaded

[a] Repeat visits in different years

[b] Data collected till end of July 2015

Table 5 Frequency Distributions of Temple Visits

	Tian Tan Buddha	Wong Tai Sin	Man Mo Temple	Tin Hau Temple	No. of visitors
Visited 3 temples	5	5	1	4	5
Visited 2 temples	30	26	8	8	36
Visited 1 temples	506	95	51	41	693
Total	541	126	60	53	734

it reflected this tourist enjoy the visit at Tian Tan so he/she revisited it. Tourist4 has visited Hong Kong in two consecutive years and visited three temples. Tourist7, 8, 29 and 30 all visited Hong Kong in two consecutive years to two temples. Both Tourist9 and 13 travelled to Hong Kong twice but there were a time gap of 4 and 3 years. Tourist9 revisited Wong Tai Sin Temple in his/her second visit where Tourist13 visited different temple in the revisit trip.

All 41 tourists have visited different temples when they travel to Hong Kong within the study period. For those who visited two temples during the same trip, Tian Tan temple was the most popular one to tourists (over 85% of the tourists visited it); and Wong Tai Sin was the second popular with 75% of the tourists took photos. In this study, only two tourists have visited three temples in one trip, and interestingly no tourist was found in the data set to visit all four temples. Besides, from the number of photos uploaded to Flickr indicated that the Asian tourists visited all temples randomly; while the tourists from Europe had rare visit to Man Mo or Tin Hau temple. For the North American tourists, Man Mo temple was the least temple they would visit.

We are aware that photo sharing behavior is varied between individuals. Some tourists might share photos of all visited temples, while others might visit two or three temples but only shared photos of one temple. For example, Tourist33 in Table 6 was found to visit two temples Tian Tan and Wong Tai Sin as indicated by the photos uploaded to Flickr. It is possible that the tourist also visited Man Mo or Tin Hau temple, but did not upload the photos. The number of visited temples should be interpreted as minimum of two temples for Tourist33. Nevertheless, the geotagged photo data still can reflect the general trend of temple visits among tourists.

4.5 Photo upload time vs. taken time

The Metadata attached with the photos consist of not only the geographic location information but also the time series information. By calculating the time difference from the date and time of photo taken to photo upload, we can work out how soon the tourists upload photos after taking. If the photos were uploaded right after

Table 6 Tourists Multiple Temple Visit Pattern by Year

	Region	Tian Tan	Wong Tai Sin	Man Mo	Tin Hau
Tourist2[a]	Asia	2010 (73), 2011 (15), 2012 (17)	–	–	2010 (6)
Tourist9[a]	Asia	–	2015 (3)	2011 (2), 2015 (3)	–
Tourist29[a]	Asia	2011 (2)	2010 (31)	–	–
Tourist30[a]	Asia	2011 (2)	–	2012 (12)	–
Tourist28	Asia	2010 (11)	–	2010 (1)	–
Tourist31	Asia	2013 (63)	–	–	2013 (10)
Tourist8[a]	HK	–	2012 (16), 2013 (17)	–	2013 (5)
Tourist32	Europe	–	2011 (1)	2011 (1)	–
Tourist33	Europe	2011 (6)	2011 (8)	–	–
Tourist34	Europe	2012 (13)	2012 (5)	–	–
Tourist35	Europe	2013 (1)	2013 (2)	–	–
Tourist36	Europe	2013 (2)	2013 (2)	–	–
Tourist37	Europe	2015 (2)	2015 (3)	–	–
Tourist38	N. America	2010 (29)	–	–	2010 (1)
Tourist39	N. America	2012 (2)	–	–	2012 (14)
Tourist40	N. America	2013 (3)	2013 (41)	–	–
Tourist41	N. America	2014 (42)	2014 (25)	–	–
Tourist1[a]	NA	2012 (17), 2014 (5)	2012 (5), 2014 (11)	–	2012 (1)
Tourist4[a]	NA	2012 (3)	2012 (2)	2011 (5)	–
Tourist7[a]	NA	2012 (3), 2013 (8)	–	2012 (1)	–
Tourist13[a]	NA	2013 (237)	2010 (67)	–	–
Tourist3	NA	2012 (56)	2012 (24)	–	2012 (3)
Tourist5	NA	2013 (191)	2013 (84)	–	2013 (8)
Tourist6	NA	2013 (2)	2013 (8)	–	2013 (3)
Tourist10	NA	2010 (4)	2010 (2)	–	–
Tourist11	NA	2010 (7)	–	–	2010 (5)
Tourist12	NA	–	2010 (7)	2010 (10)	–
Tourist14	NA	2011 (12)	2011 (5)	–	–
Tourist15	NA	2012 (1)	2012 (1)	–	–
Tourist16	NA	2012 (5)	2012 (7)	–	–
Tourist17	NA	2013 (14)	2013 (3)	–	–
Tourist18	NA	2013 (29)	2013 (1)	–	–
Tourist19	NA	2013 (8)	2013 (2)	–	–
Tourist20	NA	2014 (1)	2014 (6)	–	–
Tourist21	NA	2014 (11)	2014 (12)	–	–
Tourist22	NA	2014 (12)	2014 (4)	–	–
Tourist23	NA	2014 (9)	2014 (1)	–	–

Table 6 continued

	Region	Tian Tan	Wong Tai Sin	Man Mo	Tin Hau
Tourist24	NA	–	2014 (5)	–	2014 (3)
Tourist25	NA	–	–	2015 (4)	2015 (6)
Tourist26	NA	2015 (2)	2015 (6)	–	–
Tourist27	NA	2015 (53)	–	2015 (20)	–

Numbers in bracket indicated the number of photos uploaded

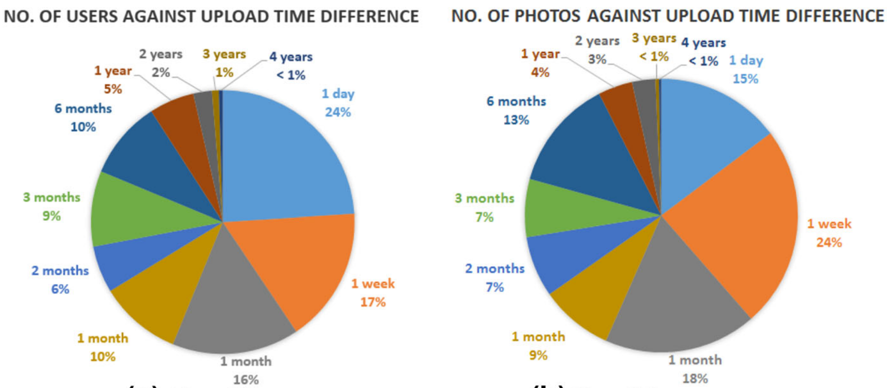
NA region information not available

[a] Repeat visits over years

Table 7 Time difference from photo taking to uploading

Time difference (uploaddate–takendate)		No. of users[a]		No. of Photos	
Less than 1 day	(<1 day)	200	21.01%	1025	14.74%
Within 1 week	(1–7 days)	138	16.57%	1658	23.84%
Within 1 month	(15–30 days)	130	15.61%	1258	18.09%
More than 1 month	(31–60 days)	84	10.08%	599	8.61%
More than 2 months	(61–90 days)	48	5.76%	502	7.22%
More than 3 months	(91–180 days)	77	9.24%	475	6.83%
More than 6 months	(181–365 days)	80	9.60%	915	13.15%
More than 1 year	(366–730 days)	46	5.52%	284	4.08%
More than 2 years	(731–1095 days)	19	2.28%	192	2.76%
More than 3 years	(1096–1460 days)	7	0.84%	29	0.42%
More than 4 years	(>1460 days)	4	1.48%	19	0.27%

[a] One visitor could have multiple temple visits on different dates



(a) No. of Users **(b)** No. of Photos

Fig. 3 Time difference from photo taking to uploading

Table 8 Ratios of photos uploaded vs. time difference

Time difference	Number of photos uploaded per visitor										
	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100	>100
Less than 1 day	0.26	0.16	0.23	0.06	0.09	0.14	0.20				0.20
Within 1 week	0.13	0.31	0.10	0.17	0.09	0.14	0.20	0.33		0.50	0.20
Within 1 month	0.16	0.17	0.37	0.36	0.15	0.43					
>1 month	0.10	0.16	0.10	0.11	0.09	0.14	0.20	0.33	0.33		
>2 months	0.06	0.01	0.07	0.08	0.18	0.14	0.20	0.33			
>3 months	0.10	0.09	0.03	0.06	0.18						0.20
>6 months	0.10	0.06	0.07	0.14					0.33	0.50	0.40
>1 year	0.06	0.03	0.03	0.03	0.18						
>2 years	0.03						0.20		0.33		
>3 years	0.01	0.01									
>4 years	0.01										

taking, the timeliness information carried by those photos was quite strong. By checking those photos, tourists at the nearby location could obtain the latest update about the travel destinations. In addition, attraction managers could able to review the context of the photo to evaluate their marketing performance from the timely photos.

4.5.1 Time difference from photo taking to uploading

Table 7 and Fig. 3 summarizes tourists uploading behaviors by counting the number of days from the date a photo was taken to the date when it was uploaded to Flickr. Nearly 80% of the photos were uploaded within 6 months. By checking the detailed photo taking and upload time slots, we found that many tourists uploaded the photos right after taking. From totally 6956 photos, nearly 15% of photos were uploaded on the same day after taking (1025 photos). Even they did not upload the photos in their trips, tourists reviewed and sort out their travel photos immediately after they were back to accommodation or home. The accumulative number of photos uploaded within 1 week was 2683 (38%), and about 40% of the photos were posted within 1 month.

Although most of the photos were uploaded to Flickr within 6 months, we still found that if a tourist took a large number of photos during the trip, he or she needed more time to sort the photos out for uploading. Normally, this task was done in 6 months to 1 year after trips, especially when the number of the uploaded photos was more than 80. The scores shown in Table 8 indicate the weights of uploaded photos against the time difference between taking and uploading. High weights in bold were found in the last three columns on the right hand side, which confirmed our findings.

4.5.2 Geographical difference

Prior study has confirmed photographs taken by individual tourists and locals were differentiated according to the duration of the period (García-Palomares et al. 2015). Our study also had shown similar findings. Table 9 shows the ANOVA results of geographical difference on the average photo upload days. It indicated there were behavioral differences among tourists from different geographical regions. Tourists from Asia, North America and Australia/New Zealand have shorter upload time. Asian and North American upload their temple photos within 1 month and Australia/New Zealand tourist upload photos in 1.5 months. European tourists generally upload photos in 3 months and South American tourists upload photos in almost 6 months.

4.5.3 Chronological difference

With the increase engagement of social media by tourists, the photos upload time shortened for the past 6 years. In 2010–2012, the average days upload were around 4–5 months. However, starting from 2013, the average days upload were no more

Table 9 Geographical Difference of Number of days uploads after Photo Taking

Region	No. of photos	Average days upload after taking			Std.	F	Sig.
		Cluster1	Cluster2	Cluster3			
Hong Kong	206	14.87[1]			78.93	42.268	0.000*
Asia	766	28.64[1]			91.58		
North America	847	32.09[1]			63.46		
Australia/New Zealand	232	47.08[1]			106.20		
Europe	885		110.19[2]		227.28		
South America	50			169.66[3]	362.86		
Total	2986						

* Significant at $p < 0.05$

[1, 2, 3] The mean difference of each region to the other two clusters are significant at the 0.05 level: $p = 0.000$

than 2 months. This has significantly shortened by at least 2 months. Table 10 shows the ANOVA result of the chronological difference for the last 6 years.

5 Conclusions, implication, limitations and future research

This study attempted to analyze tourists' visit pattern by make use of the geotagged photos uploaded by tourists to the social media sites. The proposed methodology employed geotagged photos, keyword search and P-DBSCAN clustering to identify popular attractions. The follow-up analysis indicated the different visit patterns presented by the tourists from various residential locations. Majority tourists visited one temple in their visit; only 5% visited more than one temple. Special attentions can pay on repeated tourists who visited multiple temples in one trip or the same temples at different time slots. The result can help attraction managers to identify

Table 10 Chronological Difference of Number of days uploads after Photo Taking

Year	No. of photos	Average days upload after taking			Std.	F	Sig.
		Cluster1	Cluster2	Cluster3			
2015	315	21.15[1]			24.79	89.548	0.000*
2014	1286	31.77[1]			53.94		
2013	2162	47.21[1]			95.50		
2011	1077		123.14[2]		231.08		
2010	1230		124.07[2]		300.04		
2012	886			162.49[3]	277.96		
Total	6956						

* Significant at $p < 0.05$

[1, 2, 3] The mean difference of each year to other two clusters are significant at the 0.05 level: $p = 0.000$

the repeat visitors' travel pattern. More than one-third of the photos were uploaded to social media sites within 1 week. As many tourists' spots in Hong Kong now provided free Wifi, around one-fourth of the photos were uploaded on the same day. Asian seems more attached with social media. Their average photo upload time (2 weeks) was much higher than European (4 months) and South American (5 months). The result also indicated the upload time was getting shorter from more than 4 months (in 2010) to 3 weeks (in 2015). This reflects the increasing importance of photos on social media platform. Tourists can make use of these timely photos to obtain the recent information of the attractions. The actions of photo taking and sharing do not necessarily indicate that the tourists have positive attitude in the temple. However, attraction managers and destination management organizations could attempt to use this approach to identify tourists' visit frequency and repeat visitors' behavior so as to enhance travel route design, and customize the itinerary according to the nationality of the tourists. Besides the photos and meta-data, photo comments made by photo owners or other users on Flickr can be utilize to examine the mood of photo takers toward the temples or other entities at destination in the future studies.

One of the main contributions is to propose a data extraction method for non-first tier tourist attractions, which were excluded from official reports due to time and budget constraints in the data collection using surveying approach. Attraction managers can make uses of the widely available tourists' photos from online media-sharing sites to collected relevant photos about their attractions as a proxy for the spatial distribution of tourists and understand the tourists' behavior and travel pattern. Timely photos could reflect the performance of the promotion and marketing activities. The proposed approach helps identifying popular non-first tier attractions where tourists frequently visited and shared photos, such as the case of Hong Kong temples. Pre-defined bounding boxes were setup to use GPS location of the popular attractions to accurately extract photos from social network sites. It also benefits travel agent managers, especially those who are working on identifying tourist's attractions, to examine the performance of their spots, which normally cannot obtain travel statistics from traditional research documents. The advantages of using this method to extract photos are: first, keywords only use for initial photo extraction to cluster popular tourist spots. Then all geotagged photos within the bounding box would be extracted so that even the social media users did not input any text description. As a result, no photo would be overlooked and ensure the database is completed. Furthermore, this method could help attractions managers to understand the geographical differences in photo sharing behavior among temple visitors. Moreover, they can obtain detail behavioral difference among countries. The more comprehensive data they have, their marketing strategies could be more precise.

Other social media platforms for photos sharing, such as Instagram, Twitter and Facebook, are increasing popular; and many of them provide geographical data. However, the major limitation of those social media platforms is that the original GPS location of the geotagged photos is not preserved when the photos are uploaded. Instead, the GPS location of where user posted the photos is recorded on the servers rather than the location of the photos themselves. Flickr is still one of the most popular photos sharing platforms that allow the original GPS information to be stored. This might affect the accuracy of the actual location of the photo taken if photos are extracted from other social media platforms.

This study has several limitations as well due to the nature of photos sharing on social media sites. First, this study only use Flickr as database and only 6956 photos were extracted. As this sample size was relatively small and did not cover majority of the social media sites therefore the result cannot be generalized. Second, many Flickr users did not indicate their nationality in the profile. In this study, only 43% of the users have indicated their nationality so the geographical characteristics were not generalized. Last but not least, over 70% of the inbound tourists of Hong Kong were from Mainland China, but Flickr is not popular in China so the sharing and visitor travel pattern for China market is not generalized. The number of photos upload from China may not reflect the actual tourists' interests in temples. It is worth to mention that the identification of popular temples is dependent on the photo sharing behavior of visitors. Some tourists might visit the temples in Hong Kong and took no photos or took photos but not posted on Flickr, thus they were not accounted in the collected data set. The identified temples in this study should be interpreted as temples where tourists frequently visited and shared photos on Flickr.

Future research not only can expend the number of social media sites and increase the size of data samples; the context of the photos can further be analyzed so as to understand the tourist's interests and their activities at attractions. The study of temples based on geotagged images from social media sites opens the door of exploring emerging tourist interests and visit patterns with large-scale user generated content. Future research would focus more on each individual temple and look for specific activities that tourists will be happy to participate.

Acknowledgements The work presented in this paper was partly supported by ISU Research Project Grant funded by I-Shou University, Taiwan (ISU-104-08-02A).

References

Crawshaw C, Urry J (1997) Tourism and the photographic eye. In: Rojek C, Urry J (eds) Touring cultures: transformations of travel and theory. Routledge, London, pp 176–195

Euromonitor International (2015) Top 100 city destinations ranking—analyst insight from euromonitor international. http://blog.euromonitor.com/2015/01/top-100-city-destinations-ranking.html. Accessed 21 Aug 2015

Flickr (2015) Flickr services: the app garden. https://www.flickr.com/services/api/. Accessed 18 Dec 2015

García-Palomares JC, Gutiérrez J, Mínguez C (2015) Identification of tourist hot spots based on social networks: a comparative analysis of European metropolises using photo-sharing services and GIS. Appl Geogr 63:408–417

Giannoulakis S, Tsapatsoulis N (2015) Instagram hashtags as image annotation metadata. In: Chbeir R, Manolopoulos Y, Maglogiannis I, Alhajj R (eds) Artificial intelligence applications and innovations. Springer, Switzerland, pp 206–220

González-Rodríguez MR, Martínez-Torres R, Toral S (2016) Post-visit and pre-visit tourist destination image through eWOM sentiment analysis and perceived helpfulness. Int J Contemp Hosp Manag 28(11):2609–2627

Hall MM, Jones CB, Smart P (2015) Spatial natural language generation for location description in photo captions. In: Fabrikant SI, Raubal M, Bertolotto M, Davies C, Freundschuh S, Bell S (eds) Spatial information theory. Springer, Berlin, pp 196–223

HKTB (2014) Visitor profile report 2014. http://securepartnernet.hktb.com/filemanager/intranet/dept_info/private_20/paper/Visitor-Pro/Profile2014/Visitor_Profile_2014_0.pdf. Accessed 21 Aug 2015
HKTB (2016) Visitor arrival statistics—Dec 2015. http://partnernet.hktb.com/filemanager/intranet/ViS_Stat/ViS_Stat_E/ViS_E_2015/Tourism_Statistics_12_2015_0.pdf. Accessed 21 Mar 2016
Kachkaev A, Wood J (2013) Crowd-sourced photographic content for urban recreational route planning. Paper presented at the University Transport Study Group UK Annual Conference, 2–4 Jan 2013, Oxford
Kisilevich S, Krstajic M, Keim D, Andrienko N, Andrienko G (2010) Event-based analysis of people's activities and behavior using Flickr and Panoramio geotagged photo collections. In: Information visualisation (IV), 2010 14th International Conference, IEEE, pp 289–296
Kisilevich S, Mansmann F, Keim D (2010) P-DBSCAN: a density based clustering algorithm for exploration and analysis of attractive areas using collections of geo-tagged photos. In: Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application Article. Article no. 38, Bethesda, Maryland, USA
Lau G, McKercher B (2006) Understanding tourist movement patterns in a destination: a GIS approach. Tour Hosp Res 7(1):39–49
Leung XY, Wang F, Wu B, Bai B, Stahura KA, Xie Z (2012) A social network analysis of overseas tourist movement patterns in beijing: the impact of the olympic games. Int J Tour Res 14(5):469–484
Lew A, McKercher B (2006) Modeling tourist movements: a local destination analysis. Ann Tour Res 33(2):403–423
Önder I, Koerbitz W, Hubmann-Haidvogel A (2014) Tracing tourists by their digital footprints the case of austria. J Travel Res 55(5):566–573
Pan S, Lee J, Tsai H (2014) Travel photos: motivations, image dimensions, and affective qualities of places. Tour Manag 40:59–69
Stepchenkova S, Morrison AM (2006) The destination image of Russia: from the online induced perspective. Tour Manag 27(5):943–956
Stepchenkova S, Zhan F (2013) Visual destination images of Peru: comparative content analysis of DMO and user-generated photography. Tour Manag 36:590–601
Vu HQ, Li G, Law R, Ye BH (2015) Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. Tour Manag 46:222–232
Yanai K, Kawakubo H, Qiu B (2009) A visual analysis of the relationship between word concepts and geographical locations. In: Proceedings of the ACM International Conference on Image and Video Retrieval. ACM, New York, pp 13:1–13:8
Yin Z, Cao L, Han J, Zhai C, Huang T (2011) Geographical topic discovery and comparison. Proceedings of the 20th International Conference on World Wide Web. ACM, New York, pp 247–256
Zheng YT, Zha ZJ, Chua TS (2012) Mining travel patterns from geotagged photos. ACM Trans Intell Syst Technol 3(3):1–18