

Mapping Spatiotemporal Tourist Behaviors and Hotspots Through Location-Based Photo-Sharing Service (Flickr) Data

Joey Ying Lee and Ming-Hsiang Tsou

Abstract Social media services and location-based photo-sharing applications, such as Flickr, Twitter, and Instagram, provide a promising opportunity for studying tourist behaviors and activities. Researchers can use public accessible geo-tagged photos to map and analyze hotspots and tourist activities in various tourist attractions. This research studies geo-tagged Flickr photos collected from the Grand Canyon area within 12 months (2014/12/01–2015/11/30) using kernel density estimate (KDE) mapping, Exif (Exchangeable image file format) data, and dynamic time warping (DTW) methods. Different spatiotemporal movement patterns of tourists and popular points of interests (POIs) in the Grand Canyon area are identified and visualized in GIS maps. The frequency of Flickr's monthly photos is similar (but not identical) to the actual tourist total numbers in the Grand Canyon. We found that winter tourists in the Grand Canyon explore fewer POIs comparing to summer tourists based on their Flickr data. Tourists using high-end cameras are more active and explore more POIs than tourists using smart phones photos. Weekend tourists are more likely to stay around the lodge area comparing to weekday tourists who have visited more remote areas in the park, such as the north of Pima Point. These tourist activities and spatiotemporal patterns can be used for the improvement of national park facility management, regional tourism, and local transportation plans.

Keywords Spatiotemporal • Hotspot analysis • Geo-tagged photos
Tourism • Flickr • Grand Canyon

J. Y. Lee • M.-H. Tsou (✉)

The Center for Human Dynamics in the Mobile Age, Department of Geography,
San Diego State University, San Diego, CA, USA
e-mail: mtsou@mail.sdsu.edu

J. Y. Lee

e-mail: nuo5218@gmail.com

© Springer International Publishing AG 2018

P. Kiefer et al. (eds.), *Progress in Location Based Services 2018*, Lecture Notes
in Geoinformation and Cartography, https://doi.org/10.1007/978-3-319-71470-7_16

315

1 Introduction

The tourism industry plays a key role in the economic development of many countries. Statistics from the UN's World Tourism Organization indicate that the tourism industry contributes up to 40% of the gross domestic products (GDPs) of developing countries (Ashley et al. 2007). To boost the tourism industry, the further development of existing tourism locations and identification of new tourism attractions are both recognized as crucial approaches in many countries. Tourism geography research that provides a spatial view of attractions can greatly help tourism industry development.

Over the past several years, many approaches toward mapping tourist behaviors and hotspots using photo-sharing service data have emerged (García-Palomares et al. 2015; Hawelka et al. 2014; Sun and Fan 2014; Vu et al. 2015). Conventional tourism management involves a mixed method approach using quantitative research and qualitative research based on questionnaire surveys, focus groups, and interviews. Traditional methodologies, such as questionnaires and interviews have limited capability in data collection (Chen and Chen 2012; Sun and Budruk 2015). The methodologies of tourism geography research continue to evolve as technology advances. Recently, more and more photos were taken by smartphones and GPS-equipped cameras in popular tourist attractions with geo-tagged information (Tsou 2015). Many geo-tagged photos were uploaded to photo-sharing websites such as Flickr, Instagram, and Panoramio allowing public access. Researchers can use the public application programming interfaces (APIs) to download and analyze these public accessible geo-tagged photos and analyze their spatiotemporal patterns.

This research studies geo-tagged Flickr photos collected from the Grand Canyon area within 12 months (2014/12/01–2015/11/30) using kernel density estimate (KDE) mapping, Exif (Exchangeable image file format) data, and dynamic time warping (DTW) methods. The Grand Canyon is one of the most popular tourist attractions in the U.S. and it is located across two states: Arizona and Utah (277 miles long, and up to 18 miles wide). There are over five million tourists visited the Grand Canyon during the last ten years. Natural resource management and transportation plans became important issue for the National Park Service (NPS) Agency. One key question in tourism management is to identify when and where exactly tourists are. Geo-tagging photos can indicate where and when photos have been taken by tourists, and thus can be used for tourism management.

This study utilized the space-time analysis framework for analyzing tourist behaviors and hotspots. Space-time analysis in geography was developed in the 1970s (Taaffe 1974; Palm and Pred 1974; Cullen 1972; Sauer 1974). Space-time geography can provide a comprehensive analysis framework for many research topics, such as criminology, public health, and tourism management. Space-time geography research focuses on some unique time analysis methods, such as duration, accessibility, and trajectory by using both spatial and temporal variables. Social media data can be a great data source for conducting space-time analysis

since the data include both space and time variables (Yuan and Nara 2015; Issa et al. 2017). The spatio-temporal patterns of tourists' behaviors are bound by the spatial distributions of different destinations, and they are easily affected by spatio-temporal constraints. Therefore, the patterns within the analytical construct of the space-time prism can be explored using time geography (Chen and Kwan 2012).

2 Relate Work

2.1 *Analyzing Travel Behaviors Using Geo-tagged Photos*

Researchers have developed various methods for acquiring tourist behavior data, including surveys, GPS tracking, and interviews. Since GPS devices have become inexpensive and affordable after 2000, many studies have combined GPS data with questionnaires to analyze tourist behaviors (McKercher et al. 2012). Gao et al. (2013) used the check-in social media data to perform traffic forecasting, disaster relief, and advertising services. Girardin et al. (2008a, b), Popescu and Grefenstette (2011), and Majid et al. (2013) explored the spatio-temporal pattern through Flickr photos. Girardin et al. (2008) used Flickr photos to explore the tourist behaviors. Popescu and Grefenstette (2011) used historical photos from certain Flickr users to build a personal tourist recommendation system. Majid et al. (2013) used geo-tagged Flickr photos to predict users' tourist destination preferences. Some studies use tag frequency of social media images to acquire tourist behavior patterns (Sun and Fan 2014). Therefore, analyzing social media pictures with their geo-tagged information and time stamps can be a promising method to improve tourist management and identify regional hotspots of POIs. García-Palomares et al. (2015) research identified tourist hotspots by analyzing social media data, as well as revealed the spatio-temporal patterns of the identified tourist hotspots in European cities. Furthermore, their study highlighted the difference between residents' and tourists' daily attractions and travel routes. García-Palomares et al. (2015) study relies on using spatial statistical methods (hexagons with cluster analysis) to determine tourist hotspots and using geo-tagged photos to identify tourist attractions in Barcelona. Kádár's (2014) analysis of geographically positioned photography retrieved from Flickr with tourist arrivals and registered hotel bed nights (from TourMIS website) for 16 European cities. There are high correlations between bed nights and geo-tagged Flickr images. Birenboim (2016) utilized Ecological Momentary Assessment (EMA) to conduct surveys for tourist experiences in a high resolution spatiotemporal scale. Önder et al. (2016) traced Austria tourists' travel routes using their digital footprints. In their research, they collected photos with the geo-tagged "Austria" from 2007 to 2011 using the Flickr API. To differentiate tourist photos from non-tourist ones, they used a "time span" concept where a user who uploads two different photos in two different places within a

certain period is identified as a tourist. This study also used multi-level scales to evaluate tourist footprints. The result showed that Flickr data could better represent tourism information on a city level rather than a regional level. Their research suggested that although Flickr could be used in tracking tourists' digital footprints, the accuracy of the user tracking may vary in different locations, depending on whether it is in a region level or a city level.

Aggregated geo-tagged social media can also reveal groups' semantic meanings and group activities. Kisilevich et al. (2010) used P-DBSCAN (Density-based spatial clustering of applications with noise) method to detect attractive destination from aggregated geo-tagged photos. Kennedy and Naaman (2008) used text mining to explore place semantics from Flickr tag data, while Cranshaw et al. (2012) used Foursquare data to investigate socially dynamic neighborhoods using clustered groups based on their social similarities. These research studies identified clustered groups using Flickr photos, which can reveal human mobility and to explore the patterns of human mobility.

Vu et al. (2015) used geo-tagged photos to explore travel behaviors in Hong Kong. They built a Hong Kong inbound tourist Flickr photo dataset and used a Markov chain model for travel pattern mining. Their research demonstrated that the Markov chain model could be applied to predict the probability of tourist routes between two tourist spots, and the result could be used by the government to improve transportation services. The Markov chain model could also be applied to model the tourist flows (Vu et al. 2015).

All these previous studies mentioned above did not utilize kernel density kernel density estimate (KDE) mapping nor dynamic time warping (DTW) methods for the analysis of tourist activities and hotspots, which are the major methodological contributions in this paper for tourism geography.

3 Data Collection

The data downloaded from Flickr within the Grand Canyon area in December 2014 to November 2015, which is from winter, spring, summer, and autumn, included 38,127 photos. The collection boundary of the Grand Canyon area had illustrated in Fig. 1. In this study, three types of data are acquired from Flickr APIs: time, location, and context data (Exif). The time data can be collected through timestamps of photos. The location data can be retrieved from users' instant locations via the mobile devices' coordinates or the check-in places they send along with photos (geotagged). When Flickr users upload their photos, they can choose whether they want to keep the Exif (Exchangeable image file format) info and coordinates or not. In this study, we will only collect the photo information containing coordinates and then used the photo id to retrieve their Exif information if available. The Exif data include detail information about the camera devices, such as "Manufacturer", "Model", "Date and Time

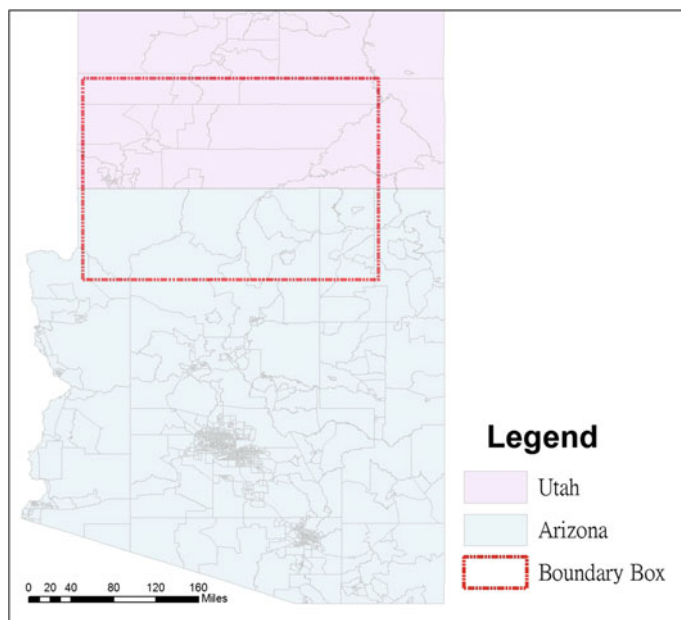


Fig. 1 The Flickr data collection boundary (the red box) across Utah and Arizona

(original)". We can use the Exif data to identify photos taken by smart phones (iPhones, Andriod phones, etc.) and cameras (Canon, Nikon, etc.).

One limitation of this study is the uncertainty of timestamps and locations. The timestamps in Flickr photos could be the time of taking photos or uploading photos. The geo-tagged locations can be modified or changed by users. Different types of spatiotemporal analysis (such as seasonal or weekend/weekday comparison) could be affected by the uncertainty of these data collection.

In this study, we used python program to collect Flickr photo information via its APIs and then stored in MongoDB database framework. The basic statistics of these photo data we collect are shown in Table 1.

As Table 1 shows, this study collected 38127 photo information in 2014/12/01–2015/11/30. In which 25395 were collected with coordinates (geo-tagged) in 2015. Among these geo-tagged photo information collected in 2015, 7471 (29.4%) of photos were taken in weekends, 17924 (70.6%) of photos were taken in weekdays. For the monthly change, May is the highest month for Flickr photo uploaded count, the winter months, which is from December to February, are the lowest month.

As Fig. 2, among these photos taken by the camera, Canon and Nikon were the most popular camera devices. 353 photos were taken by *Canon EOS 6D*, 207 photos were taken by *Nikon D600*, 123 photos were taken by *Canon EOS 7D*, and 113 photos were *Nikon D7100*. These cameras are all digital single-lens reflex camera (DSLR), which means the users are likely professional users. As Fig. 3, among those photos taken by phones, iPhone was the most popular device.

Table 1 Descriptive analysis table for photo collection in the Grand Canyon area

Total photo collected in 2015 (2014/12/01–2015/11/30)	38127	Total photo collected with coordinates (geo-tagged) in 2015	25395	(66%)	
	Count by different days	Count by devices	Count by seasons		
	Weekend	7471 (29.4%)	Non-specific	23660 (93%)	Winter 2625 (10.3%)
	Weekday	17924 (70.6%)	Cameras	1377 (5.4%)	Spring 10056 (39.6%)
			Smart phones	358 (1.6%)	Summer 7105 (28.0%)
					Autumn 5609 (22.1%)
Total		25395		25395	25395

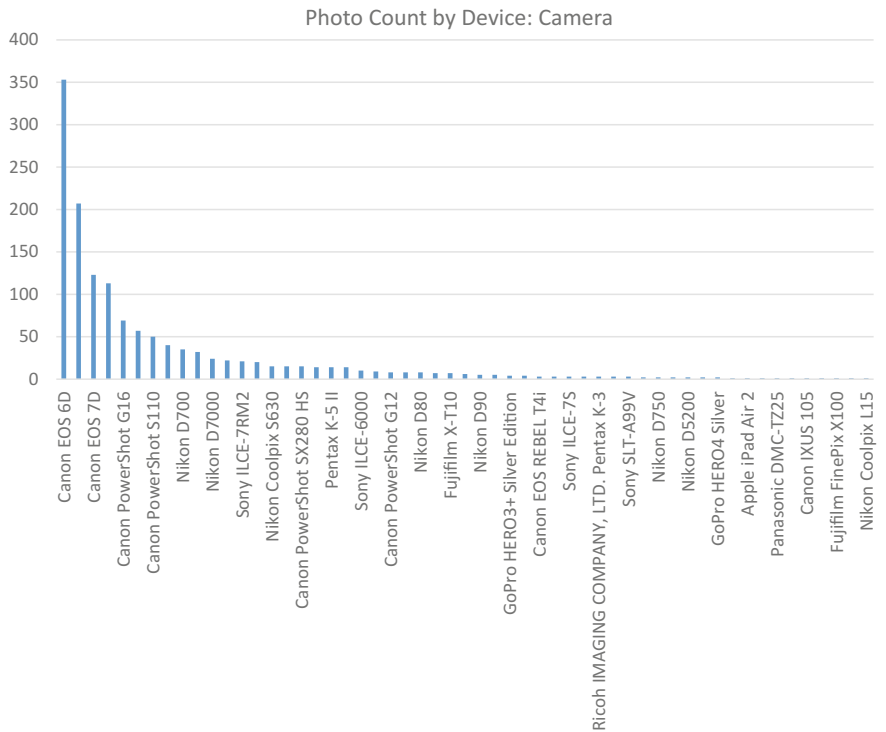


Fig. 2 Photos taken counts by different camera devices

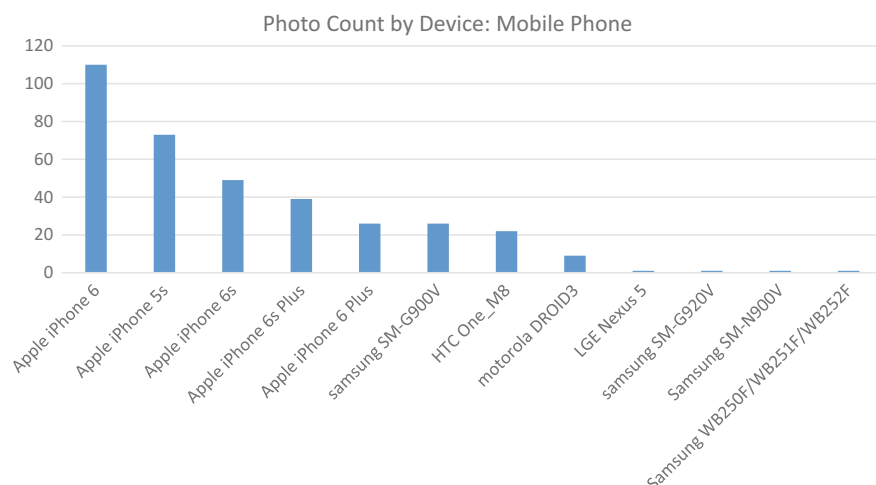


Fig. 3 Photos taken counts by different smart phone models

110 photos were taken by iPhone 6 and 73 photos were taken by iPhone 5s. 49 photos were taken by iPhone 6s and 39 photos were iPhone 6s Plus. 26 photos were taken by iPhone 6 plus.

4 Research Method

4.1 Kernel Density Estimation Mapping

To analyze the statistical outcome and identify hotspots of tourist behaviors, Kernel Density Estimation (KDE) mapping has been implemented in this research. KDE mapping is able to identify the hotspots visually from large datasets (Okabe et al. 2009; Tsou et al. 2013a). Using KDE has been widely used to create raster files to explore hotspots in social media research. Han et al. (2015) used KDE to identify hotspots using Twitter data and by exploring Twitter activity. Han also established the differential maps to compare the changes in activity by using the raster-based “map algebra tool” developed by ESRI after KDE hotspot maps were created. Following the method used by Tsou et al. (2013b), the below formula was applied to the raster formatted maps for all case studies.

Differential Map = (Each Cell Value of Map A/Maximum Cell Value of Map A) – (Each Cell Value of Map B/Maximum Cell Value of Map B)

One important variable in the KDE method is the kernel radius. Adopting different sizes of kernel radius will generate different scale of hotspot analysis. This study utilized two spatial scales of KDE for tourist activity analysis. The first level

is 50 km which can be used to identify the general (large regions) hotspots in the Grand Canyon area. The second level is 200 m which can identify smaller hotspots along with roads and trails (with a higher spatial resolution).

4.2 *Dynamic Time Warping*

To compare the similarities and difference among the trajectories, distance is used as the common variable (Tan et al. 2005), which means if the distance is lower, the similarity is higher. Dynamic Time Warping (DTW) is one of the methods used with distance to compare the trajectories. In this study, the Flickr API will be employed to generate data with user IDs, as well as coordination and time data. However, while these photos were located as point data the tourist trajectory for each user will still present a crucial problem that must be solved. Therefore, Python was used to write a module to solve this problem. DTW distance value is a comparison value between two users. In this research, the DTW distance value of two users was calculated by Python. When the distance is lower, the similarity is higher.

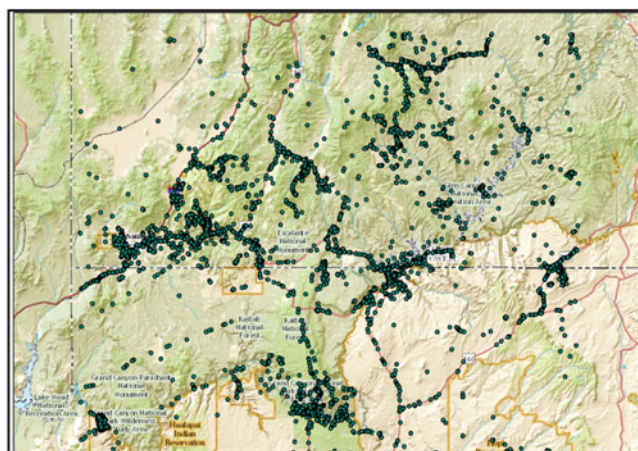
5 Major Findings

Two types of tourist activity analysis were conducted by using the 2015 geo-tagged Flickr data in the Grand Canyon area: spatiotemporal hotspot analysis (with two case studies), and tourist trajectory analysis. Figure 4 illustrated the spatial distribution of geotagged photos (top) in the Grand Canyon area and the activity hotspot map (bottom) using 50 km radius KDE method. The hotspot map illustrated two popular tourist locations within the study area: Grand Canyon Village area (Visitor Center) and Emerald Pools. Therefore, we selected the two sub-regions as our case studies.

5.1 *Spatiotemporal Hotspot Analysis*

The first case study area is the Grand Canyon (GC) Village (Visitor Center) area, which is not far from the south entrance of the park. GC Village would usually be the first stop for visitors. It provides all kinds of facilities like hotels, visitor center, restaurants, and gift shops (Fig. 5). Figure 5 illustrated some hotspots of geo-tagged Flickr photos using 200 m radius KDE method. The hotspots (red color) are near Grant Canyon Village, Hopi House and some scenery locations.

The second case study area, Emerald Pools (Fig. 6), in Zion was in the heart of Zion Canyon, near Zion Lodge. It has a variety of accommodations and a dining



Activity Hotspots in the Grand Canyon Area

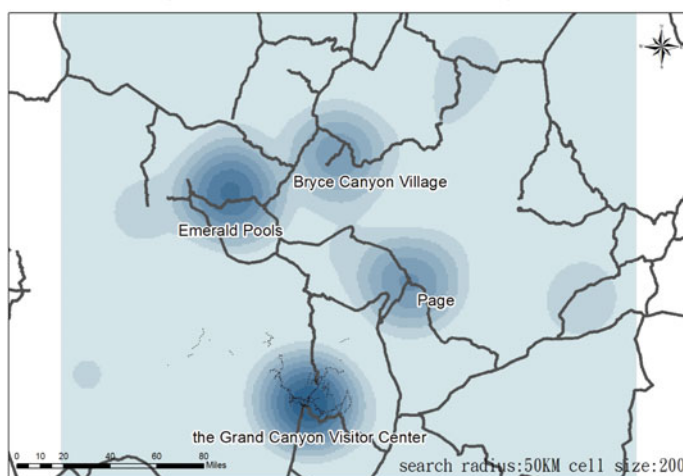


Fig. 4 The Flickr geotagged photos (top) and the activity hotspot map (bottom) using 50 km radius KDE method

room for visitors. To the west of Zion Lodge locates Lady Mountain, which is one of Zion's landmarks.

To explore how these hotspots changed pattern, differential map has applied on the two case studies. There are three major colors used throughout the two differential maps: (1) blue, (2) green, and (3) red. The blue areas show decreased photo activity density, the green areas show constant density, and the red areas show increased photo activity density.

In the Grand Canyon visitor center area, the photos taken in summer are dispersed off the trails and broader regions comparing to winter (Fig. 7). In case study 2 (Fig. 8), Emerald Pools, the photos taken in summer also shows similar patterns

Kernel Density Estimation Map in Grand Canyon Visitor Center

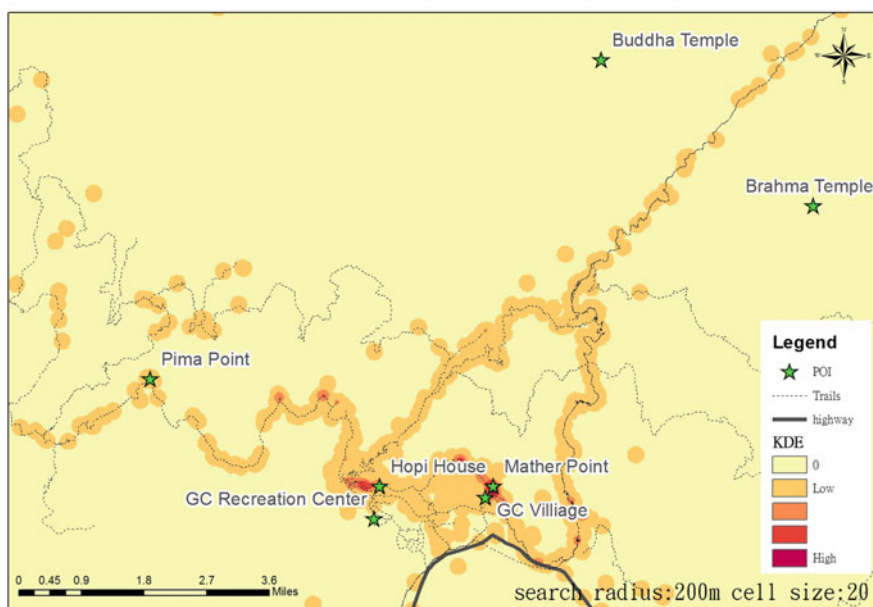


Fig. 5 The Flickr photos hotspots (red) in the Grand Canyon Visitor Center and Village Areas using 200 m radius KDE

(more trails). Most photos taken in winter are clustering within Zion Lodge, Echo Canyon Passage, and Zion Observation Point. On the northwest of Fig. 8 map, it revealed only two photos taken in winter, but many photos were taken in summer. For the seasonal patterns revealed in Fig. 8, it showed that the tourist activity areas have been influenced by season.

Not only the seasonal patterns can be revealed through Flickr data, weekday and weekend patterns also can be explored. For exploring weekday and weekend patterns in the Grand Canyon area, Fig. 9 used differential mapping to identify the differences between weekdays and weekends. Photos taken in weekdays are more disperse in the park, such as the north of Pima Point comparing photos taken in weekends. Not only the spatial difference of their disperse but also the average travel times spent in the Grand Canyon area are different between weekdays and weekends (Table 2). The weekday travel duration is 68 h, and the weekend travel duration is only 45 h. The other case study also showed similar patterns (Fig. 10).

In this research, the Exif information of each photo had been collected, per the statistics, most of the camera users use DSLR to take the photos, and most of the phone users used iPhone. In different devices, such as camera and phone, tourist behavior may be different. Although the three case study areas are not in the same area, the similar spatio-temporal patterns still can be identified. The camera users can travel apart from main tourist area, even choose the unpaved trail. The smart

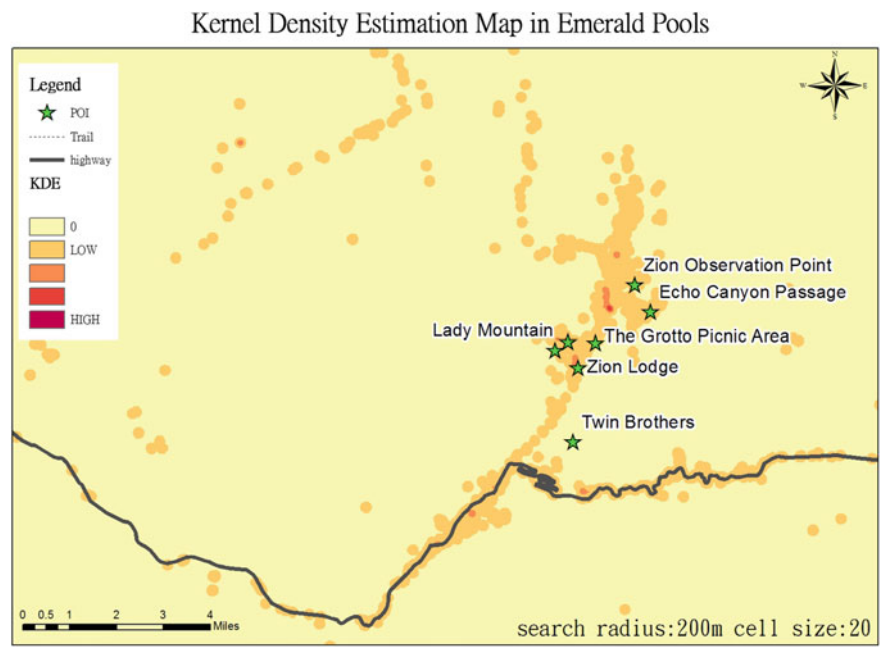


Fig. 6 The Flickr photos hotspots in the Emerald Pools Areas

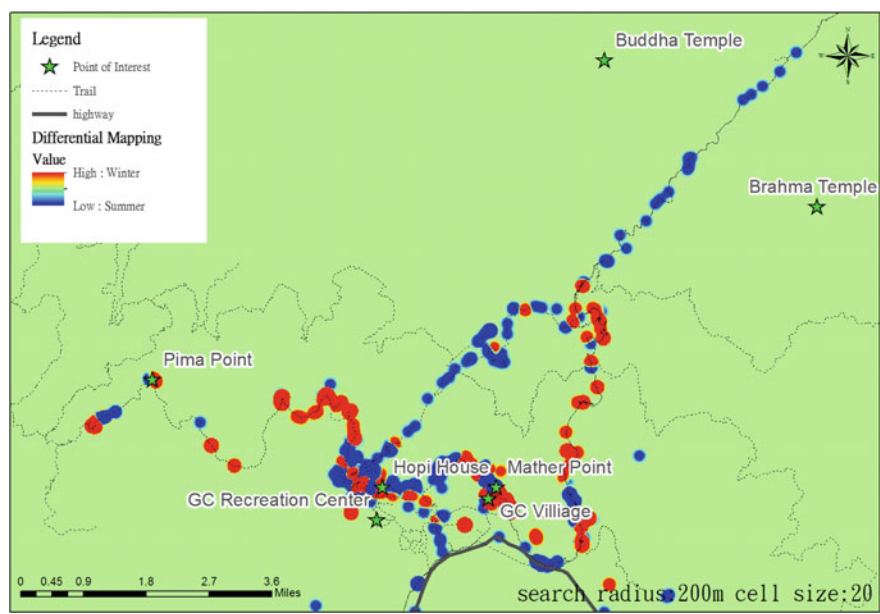


Fig. 7 The Flickr photo differential map for the Grand Canyon Visitor Center case study: using seasonal difference (Blue: Summer, Red: Winter)

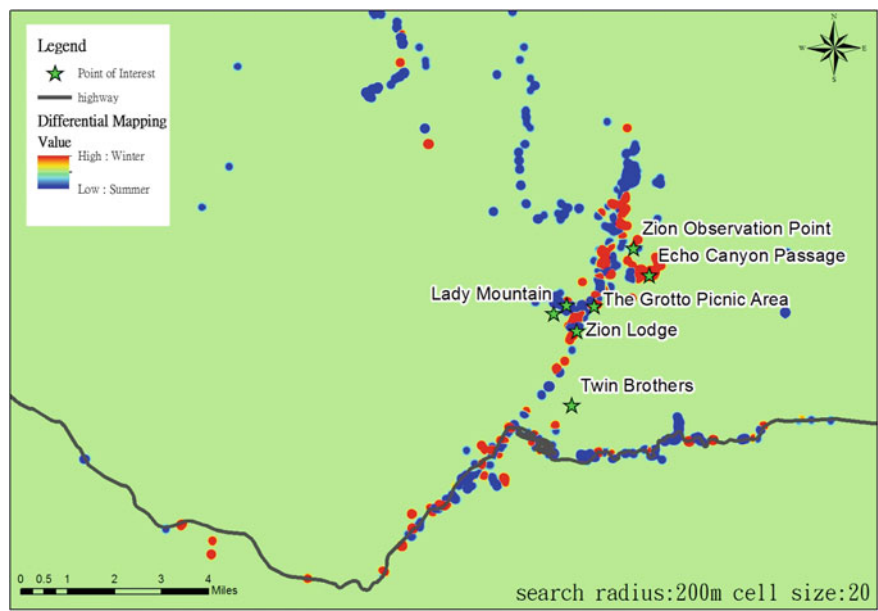


Fig. 8 The Flickr photo differential map for the Emerald pools case study: using seasonal difference (Blue: Summer, Red: Winter)

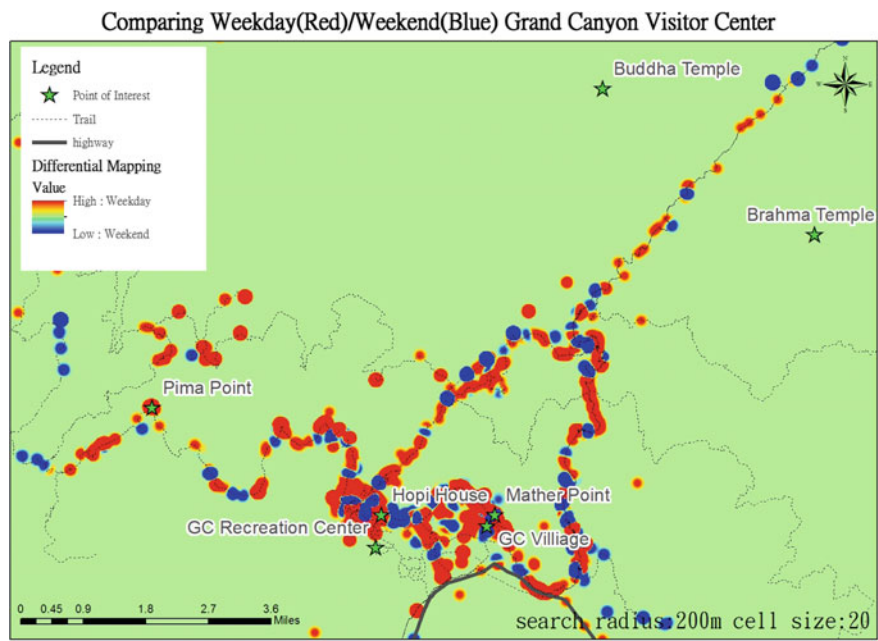


Fig. 9 The Flickr photo differential map for the Grand Canyon Visitor Center case study: Comparing weekday (blue) and weekend (red) difference

Table 2 The average travel times on weekdays and weekends

	Average travel duration (h)
Weekday travels	68
Weekend travels	45
All	63

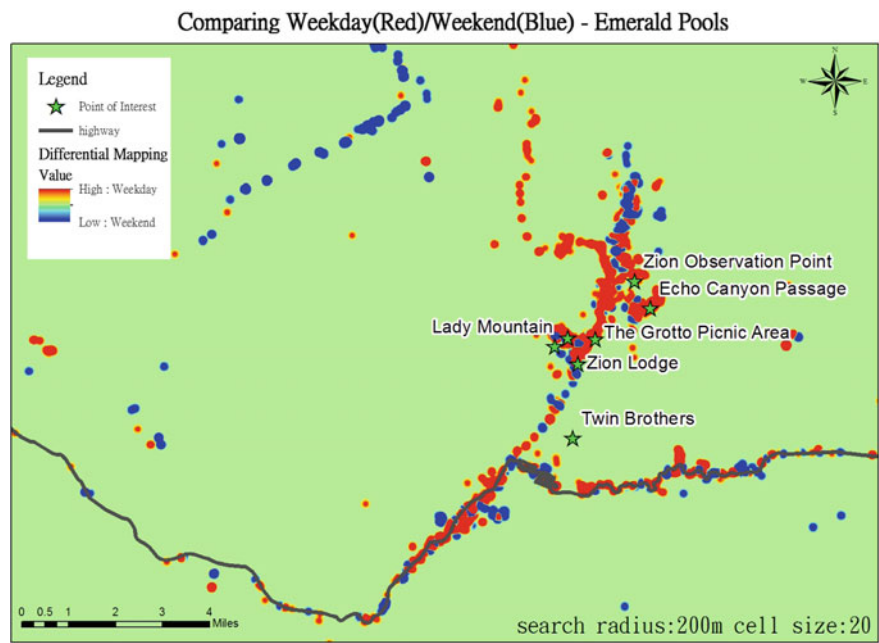


Fig. 10 The Flickr photo differential map for the Emerald Pools case study: Comparing weekday (blue) and weekend (red) difference

phone users usually travel in the recommended locations suggests by National Park Service. In the Grand Canyon Village, the camera users’ footprint could be found at the “unpaved trail” in the west of Pima Point, which is different with smart phone users (Fig. 11). In Emerald Pools, it can be found the photos taken by phone are more likely cluster near Zion observation point and The Grotto Picnic area, which is the tourist destination suggested by national park service (Fig. 12).

5.2 Travel Time and Dynamic Time Warping for Trajectory Analysis

Geo-tagged photographs on Flickr platform showed many photos taken by tourists or local residents. The criterion for determining whether the photographs were

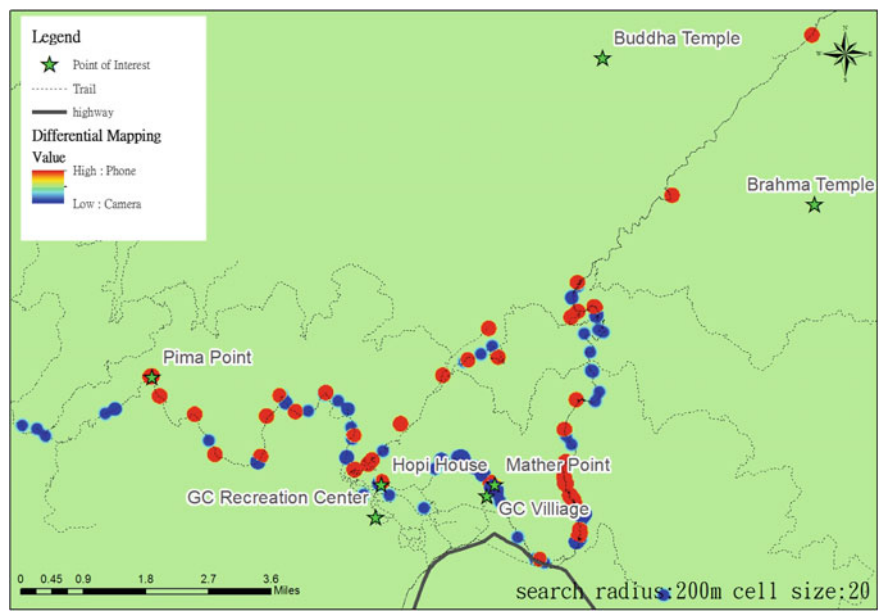


Fig. 11 The Flickr photo differential map for the Grand Canyon Visitor Center case study: using the different devices: smart phones (Red) and cameras (Blue)

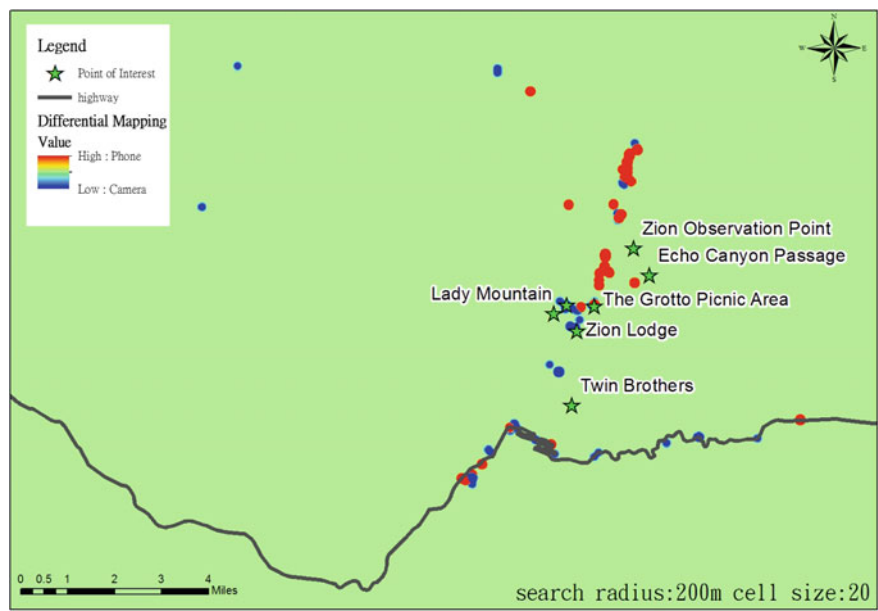


Fig. 12 The Flickr photo differential map for the Emerald pools case study: using the different device: phone (red) and camera (blue)

Table 3 The average travel time for visitor in different season

Total Flickr user visitors (estimated)	735	Total visitor average travel duration	63.53 h
winter	108	Average travel duration	49.03 h
spring	272	Average travel duration	68.81 h
summer	221	Average travel duration	59.00 h
autumn	134	Average travel duration	70.56 h

taken by visitors or local residents in this study is the time period during which each user had taken pictures: if this period exceeded one month, then the photographs were classified as taken by residents; if the period was less than one month (720 h), then the users were classified as tourists. Table 3 illustrated the estimated temporal patterns of visitors in Grant Canyon area and their average travel duration.

Selected Camera Users

To analyze visitors' trajectory patterns, we selected the top three photo-uploaded users who took the highest numbers of photos by cameras as our case studies. Although these top users can not represent the average visitor's movement patterns in the Grant Canyon area, we used these cases to demonstrate the feasibility of DTW for trajectory analysis. To protect users' privacy, we used "Camera user A", "Camera user B", and "Camera user C" to label these top users. "Camera User A" joined Flickr in April 2013, and the user indicated that *"Not a regular user, just wanted to share my photos to give back to the community for some of the great photos I've seen here."* "Camera User B" joined Flickr in September 2009. The user B uploaded 11714 photos on Flickr platform, and 10700 photos are geo-tagged. "Camera user C" is similar to user B, he joined Flickr in August 2007, and the user has 7600 geo-tagged photos of total 9177 photos uploaded on Flickr platform (Fig. 13). To compare with the similarity of each user's trajectory, we used DTW analysis to measure their similarity in the distance. The DTW value of Camera users A, B, and C are on Table 4, per their DTW distance, camera user A is 194, which is higher than average. For camera user B and C, they have similar DTW distance value and similar routing per their trajectories.

Selected Smart Phone Users

For the smart phone user group, we selected the top three photo-uploaded users who took the highest numbers of photos by Smart Phones. Users who used the phone to take the photos will be defined as "leisure tourist." To protect their privacy, we used "Phone user D", "Phone user E," and "Phone user F" to label these users. "Phone user D" joined Flickr in September 2013. The user D uploaded 241 geo-tagged photos of total 356 photos. "Phone user E" joined Flickr in May 2008, the user uploaded 539 geo-tagged photos of total 4059 photos. User E identified himself/herself as a Montana resident, like hiking, backpacking, car camping, road and trail running, cross-country skiing, snowshoeing and travel.

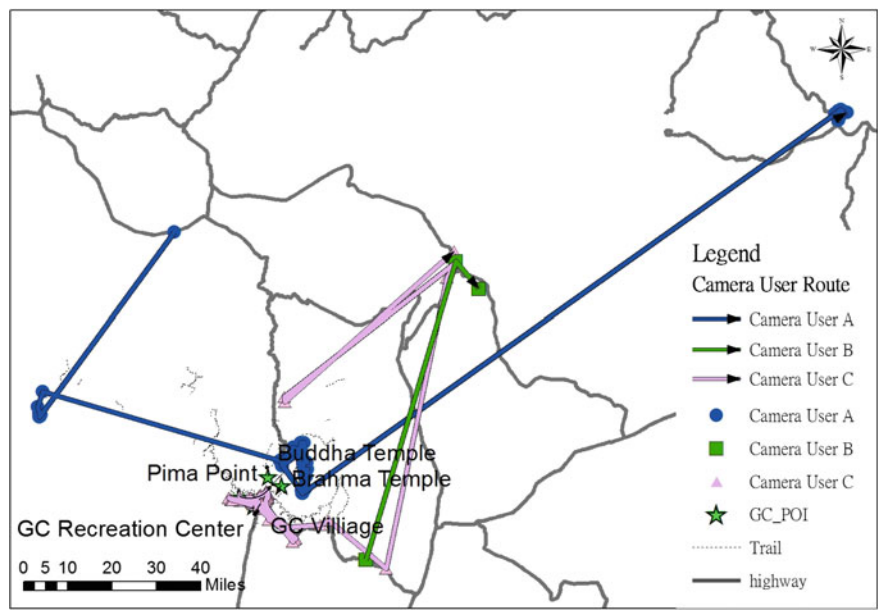


Fig. 13 Selected three camera users’ routes (A, B, and C)

Table 4 Camera users’ DTW distance value (User B and C are more similar)

	DTW distance average value (compared to other 734 routes)	Camera User A route	Camera User B route	Camera User C route
Camera User A route	194 (less representative)	–	160	99
Camera User B route	121	160	–	56
Camera User C route	77 (more representative)	99	56	–

“Phone user F” joined Flickr in April 2010. User-F had uploaded 901 geo-tagged photos, and 768 of them are public. To compare with the similarity of each user’s trajectory, the DTW distance value of phone users D, E and F had been shown in Table 5. Phone users are more likely travel within the area suggest by National Park Service, and their travel destination is more similar. Such as phone user D and F, their DTW value is only 13, and their trajectory has overlay within Emerald pools area (Fig. 14).

Table 5 Selected phone users DTW value (User D route and F route are very similar)

	DTW distance average value (compared to other 734 routes)	Phone User D route	Phone User E route	Phone User F route
Phone User D route	61	–	69	13 (very similar routes)
Phone User E route	60	69	–	199
Phone User F route	41	13 (very similar routes)	199	–

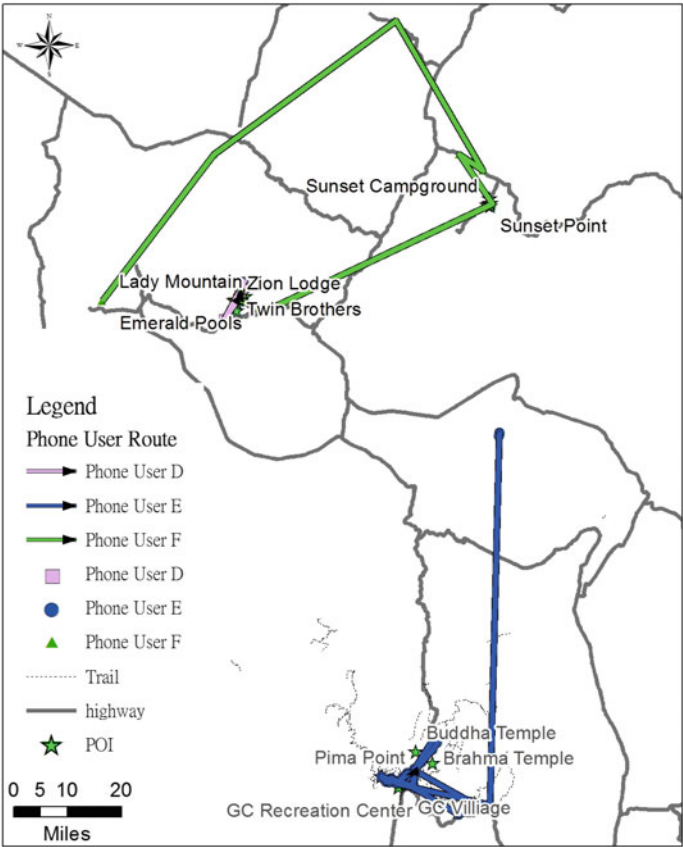


Fig. 14 Selected three smart phone users' routes (D, E, F)

This study only selected top three contributors of camera and smart phone users to demonstrate the feasibility of using DTW to compare the similarity of visitors' trajectories in the Grand Canyon area. By calculating the DTW distance value, we can find out which user may have more similar trajectory movements comparing to all other users. The strength of DTW analysis is to provide a quantitative value to compare the similarity of visitors' trajectory movement. The weakness is the missing of spatial factors and location-based analysis in the DTW analysis.

6 Conclusion

For tourism study, social media data is a new world to explore. In the past, data collection was expensive, monopolized. With social media data, researchers could collect high-resolution spatiotemporal data from public social media APIs and analyze tourist activities and behaviors. In this study, we collected and cleaned the geo-tagged Flickr photo data, and then applied two spatio-temporal analysis methods (KDE and DTW) to explore the tourists' spatial and temporal activity patterns.

The major scientific contribution in this research is to demonstrate the feasibility of using kernel density estimate (KDE) mapping for tourism hotspot analysis and dynamic time warping (DTW) methods for visitor's trajectory analysis. Previous tourism geography research works mentioned in Sect. 2 (literature review) did not utilize any kernel density methods nor DTW methods. This research also illustrated that adopting different sizes of kernel radius will generate different scale of hotspot analysis. This study utilized two spatial scales of KDE for tourist activity analysis. The first level is 50 km which can be used to identify the general (large regions) tourism hotspots in the Grand Canyon area. The second level is 200 m which can identify smaller hotspots along with roads and trails (with a higher spatial resolution).

This research identified unique activity patterns between different types of users on Flickr: camera users are exploring remote areas beyond traditional tourist attractions. Smart phone users are more likely clustered within the lodge area and viewpoints suggested by the tour guides. For temporal pattern analysis, this research identified weekday tourists are more "activate" comparing to weekend visitors. In the Grand Canyon areas, Flickr photo data can also identify the seasonal pattern: the winter photo amount is the lowest and the increased trend for spring and summer.

There are several limitations and challenges in this study. First, the demographics of Flickr photo users might be biased comparing to the general visitor profiles in the Grand Canyon area. User privacy concerns and restriction are an important issue for using Flickr data. Although this study only collected public accessible Flickr photos, the detailed trajectory analysis might reveal some personal information regarding specific users. Finally, the user's travel trajectories may not

reflect the most reality tourists' trajectories since we only collect the top 3 most active users.

Two future research directions could be explored in Flickr-based tourism research: computer image processing and text analysis. Computer image process technology using machine learning tools and deep learning methods could be used to identify the content of photos in the Grand Canyon area to explore the activity type in each photo. Text analysis, such as topic modeling or latent dirichlet allocation (LDA) methods, can be used to aggregate the texts and tags associated with each photos and provide additional information for various analysis, such as emotional analysis, social network analysis, and user profile analysis.

Acknowledgements This material is based upon work supported by the National Science Foundation, under Grant No. 1416509 and Grant No. 163464. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

- Ashley C, De Brine P, Lehr A, Wilde H (2007) The role of the tourism sector in expanding economic opportunity. John F. Kennedy School of Government, Harvard University, Cambridge
- Birenboim A (2016) New approaches to the study of tourist experiences in time and space. *Tour Geogr* 18(1):9–17
- Chen CF, Chen PC (2012) Research note: exploring tourists' stated preferences for heritage tourism services—the case of Tainan city, Taiwan. *Tour Econ* 18(2):457–464
- Chen X, Kwan MP (2012) Choice set formation with multiple flexible activities under space–time constraints. *Int J Geogr Inf Sci* 26(5):941–961
- Cranshaw J, Schwartz R, Hong JI, Sadeh N (2012) The livelihoods project: utilizing social media to understand the dynamics of a city
- Cullen IG (1972) Space, time and the disruption of behaviour in cities. *Environ Plann A* 4(4):459–470
- Gao H et al (2013) Exploring temporal effects for location recommendation on location-based social networks. In: *Proceedings of the 7th ACM conference on recommender systems*
- 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
- Girardin F, Calabrese F, Dal Fiore F, Ratti C, Blat J (2008a) Digital footprinting: uncovering tourists with user-generated content. *IEEE Pervasive Comput* 7(4)
- Girardin F et al (2008b) Digital footprinting: uncovering tourists with user-generated content. *IEEE Pervasive Comput* 7(4):36–43
- Han SY, Tsou MH, Clarke KC (2015) Do global cities enable global views? Using Twitter to quantify the level of geographical awareness of U.S. cities. *PLoS ONE* 10(7):e0132464. <https://doi.org/10.1371/journal.pone.0132464>
- Hawelka B, Sitko I, Beinart E, Sobolevsky S, Kazakopoulos P, Ratti C (2014) Geo-located Twitter as proxy for global mobility patterns. *Cartogr Geogr Inf Sci* 41(3):260–271
- Issa E, Tsou MH, Nara A, Spitzberg B (2017) Understanding the spatio-temporal characteristics of Twitter data with geotagged and nongeotagged content: two case studies with the topic of flu and Ted (movie). *Ann GIS* 23:219–235

- Kádár B (2014) Measuring tourist activities in cities using geotagged photography. *Tour Geogr* 16 (1):88–104
- Kennedy LS, Naaman M (2008) Generating diverse and representative image search results for landmarks. In: *Proceedings of the 17th international conference on world wide web*. ACM, pp 297–306
- 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 and application*. ACM, p 38
- Majid A et al (2013) A context-aware personalized travel recommendation system based on geotagged social media data mining. *Int J Geogr Inf Sci* 27(4):662–684
- McKercher B, Shoval N, Ng E, Birenboim A (2012) First and repeat visitor behaviour: GPS tracking and GIS analysis in Hong Kong. *Tour Geogr* 14(1):147–161
- Okabe A, Satoh T, Sugihara K (2009) A kernel density estimation method for networks, its computational method and a GIS-based tool. *Int J Geogr Inf Sci* 23(1):7–32
- Önder I, Koerbitz W, Hubmann-Haidvogel A (2016) Tracing tourists by their digital footprints: the case of Austria. *J Travel Res* 55(5):566–573
- Palm R, Pred AR (1974) A time-geographic perspective on problems of inequality for women (No. 236). Institute of Urban & Regional Development, University of California
- Popescu A, Grefenstette G (2011) Mining social media to create personalized recommendations for tourist visits. In: *Proceedings of the 2nd international conference on computing for geospatial research and applications*
- Sauer CO (1974) The fourth dimension of geography. *Ann Assoc Am Geogr* 64(2):189–192
- Sun YY, Budruk M (2015) The moderating effect of nationality on crowding perception, its antecedents, and coping behaviours: a study of an urban heritage site in Taiwan. *Curr Issues Tour* 1–19
- Sun Y, Fan H (2014) Event identification from georeferenced images. In: *Connecting a digital Europe through location and place*. Springer International Publishing, pp 73–88
- Taafe EJ (1974) The spatial view in context. *Ann Assoc Am Geogr* 64(1):1–16
- Tan PN, Steinbach M, Kumar V (2005). *Introduction to data mining*, 1st edn
- Tsou MH (2015) Research challenges and opportunities in mapping social media and big data. *Cartogr Geogr Inf Sci* 42(sup1):70–74
- Tsou MH, Kim IH, Wandersee S, Lusher D, An L, Spitzberg B, Gupta D, Gawron JM, Smith J, Yang JA, Han SY (2013a) Mapping ideas from cyberspace to realspace: visualizing the spatial context of keywords from web page search results. *Int J Digit Earth* 7:4. <https://doi.org/10.1080/17538947.2013.781240>
- Tsou MH, Yang JA, Lusher D, Han SY, Spitzberg B, Gawron JM, Gupta D, An L (2013b) Mapping social activities and concepts with social media (Twitter) and web search engines (Yahoo and Bing): a case study in 2012 US Presidential Election. *Cartogr Geogr Inf Sci* 40 (4):337–348. <https://doi.org/10.1080/15230406.2013.799738>
- 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
- Yuan M, Nara A (2015) Space-time analytics of tracks for the understanding of patterns of life. In: *Space-time integration in geography and GIScience*. Springer Netherlands, pp 373–398