Strategic Sustainable Development Through the Lenses of Tourists’ Preferences: A Geotagged Photo Approach

Zhihua Zhang¹, Rachel J. C. Chen² & Lee D. Han¹

¹ Department of Civil and Environmental Engineering, the University of Tennessee, USA
² Department of Retail, Hospitality, and Tourism Management, the University of Tennessee, Knoxville, USA

Correspondence: Rachel J. C. Chen, Department of Retail, Hospitality, and Tourism Management, the University of Tennessee, 246 Jessie Harris Building, Knoxville, TN 37996-1911, USA. Tel: 1-865-974-0505. E-Mail: rchen@utk.edu

Received: March 22, 2019      Accepted: April 16, 2019      Online Published: April 26, 2019

doi:10.5539/jsd.v12n3p12                  URL: https://doi.org/10.5539/jsd.v12n3p12

Abstract

Knowing tourists’ preferences and experiences with respect to their national park visits is of great importance to implementing strategically sustainable development of national parks. Flickr geotagged photos are utilized and analyzed as our main data source. We included 7090 photos from 626 people in the Great Smoky Mountains National Park to investigate visitors’ behaviors through the presentations of photo spatial and temporal patterns. The results indicated that tourist behaviors that reflect on what they like and enjoy during their park visits can be extracted from geotagged social media data in terms of frequency and length of enjoyment as visitors’ preferred spots.

Keywords: tourists’ preferences, geotagged photo, sustainable development

1. Introduction

According to the National Park Service (NPS), fifty-nine national parks in the United States received over 80 million recreational visits in 2017 (NPS, 2016). The flourishing tourism businesses of national parks have contributed significant and continuous economic benefits to local communities (Chen, Bloomfield, & Fu, 2003). The heavy traffic congestion caused by visitors have generated challenges to the effectiveness of national park management such as monitoring carrying capacity, mandating environmental conservation, maintaining a positive satisfaction level for tourists, and sustaining quality of life of the residents. To ensure the sustainable development of national park tourism, it is of great value and importance to have an in-depth and broad understanding of tourist preferences and behaviors in national parks.

Because of the proliferation of social media and the ubiquity of portable devices like smartphones, large geotagged datasets (e.g., geotagged texts, geotagged photos, geotagged videos, etc.) have become increasingly recognized over the past decade (Chua, Servillo, Marcheggiani, & Moere, 2016). These datasets can be used to extract spatiotemporal information, providing opportunities and directions for further concurrent tourism research. For example, photo-sharing websites (e.g., Flickr, Instagram, Picasa) allow users to share photos taken in places they visited embedded with time and geographical information, which can then be analyzed to project tourist behaviors and preferences. Numerous studies have been conducted and integrated with various approaches through geotagged photos (Donaire, Camprubi, & Gali, 2014; Li, Goodchild, & Xu, 2013; Majid, Chen, Mirza, Hussain, & Chen, 2015; Önder, Koerbitz, & Hubmann-Haidvogel, 2016; Sun, Fan, Bakillah, & Zipf, 2015; Vu, Li, Law, & Ye, 2015). The majority of these geotagged photo-related studies were previously conducted at a city, state, or country level, rather than at a scenic spot such as a specific national park.

The aim of this article is to investigate the spatiotemporal characteristics of tourists’ preferences and experiences in national parks using geotagged photos. Photos marked with time and geolocation information on Flickr are chosen as the main dataset because they are easily accessible. The Great Smoky Mountains National Park (GSMNP) is selected as the studied destination because it is the most-visited national park, with an average of 10 million visitors annually in the USA receiving an average of 10 million visitors annually around the world. National park managers and policymakers are facing challenges and opportunities that could change the future of the park system and lead to sustainable development through innovative strategies and technologies. For instance,
Flickr is able to offer opportunities and possibilities in terms of understanding how people use and respond to nature in national parks in a timely manner (real-time and onsite behaviors). Using the real-time geotagged photo approach, the outcomes could be compared with traditional tourist information collecting methods such as surveys or interviews for future research.

2. Literature Review

2.1 National Park and Tourism

National parks are not only frequently reserved for protect unique, or diverse natural and cultural features, but also are considered as perfect places for recreation (Lai, Hsu, & Nepal, 2013). Since the first national park, Yellowstone National Park, was established in 1872 in the United States, enormous amount of national parks have been established around the world (Reinius & Fredman, 2007). By 2016, 59 national parks are existed in the United States with over 80 million recreation visits around the world. Generally, the recreation visits to US national parks are showing an increase trend since 1979, indicating the popularity of national parks as tourist attractions.

However, the relationship between tourism and national parks is complex owing to the economic distinct economic focus of tourism and the contrasting conservation focus of national parks (Whitelaw, King, & Tolkach, 2014). Although tourism provides economic benefits for local community and tourism managers, which may facilitate the national parks conservation (Wilson, Nielsen, & Buultjens, 2009), serval negative effects are also associated with the huge volume of visitation. To minimize the environmental lost and maximize economic benefit, numerous studies have been exploring this issue, and found that a balance between nature and environment, on one hand, and tourism and recreation, on the other, is of great importance to be achieved (Beunen, Regnerus, & Jaarsma, 2008; Buckley, 2003; Eagles et al., 2013; Jamal & Stronza, 2009; Wilson et al., 2009). Thus, challenges or pressures are put on national park managers, and innovative means to manage the national parks are needed. Fortunately, with the advent of social media like Flickr, new opportunities and possibilities are created to understand how people use and respond to nature in national parks, comparing with traditional tourist information collecting method, such as surveys or interviews.

2.2 Geotagged Photos Research

The widespread use of social media has provided more possibilities for tourism research because of the ease of sharing datasets and thoughts through social media channels. Several studies reported how and why tourists posted and shared their geotagged photos (Donaire et al., 2014; Levin, Lechner, & Brown, 2017; Majid et al., 2015; Mayer, Müller, Woltering, Arnegger, & Job, 2010; Önder, 2017; Reimann, Lamp, & Palang, 2011). Among these studies, three main fruitful lines are identified and classified including preference identification, visitors’ onsite movement, and travel recommendations. The first line is tourist attractions and preferences identification. Based on the assumption that people tend to share the pictures from where they like most, some studies have been conducted to explore tourist preferences and identify tourist attractions. For example, with the photographs uploaded on Flickr from 183 cities of different size around the world, Popescu, Grefenstette, and Moëllic (2009) developed a mining approach to explore the places people visited, the duration of their stay and panoramic spots. García-Palomares, Gutiérrez, and Mínguez (2015) identified and analyzed the main tourist attractions in eight major European cities (Athens, Barcelona, Berlin, London, Madrid, Paris, Rome, and Rotterdam) and performed a comparison analysis between cities with geotagged photographs on Panoramio. Meanwhile, scholars have investigated the tourist preferences by exploring the spatial and temporal patterns of visitors’ geotagged photos (Bae & Yun, 2016; Li et al., 2013; Su, Wan, Hu, & Cai, 2016). For example, Li et al. (2013) explored spatiotemporal footprints of visitors with georeferenced data generated in Flickr and Twitter within the contiguous United States, and deeply investigated the patterns with the characteristics of visitors in different counties of California.

Mining tourist movement is another research theme for geotagged photo research. With the spatiotemporal metadata of geotagged photos, researchers are able to trace visitors’ footprints. To understand the visitors’ movements, including where they are going, departure time to their destination, and where they are likely to go next, Bermingham and Lee (2014) presented a novel spatiotemporal sequential pattern mining approach and tested the effectiveness of their analytical approach using a large dataset of Flickr photos in Queensland in 2012. To capture the movement trajectories of tourists, Vu et al. (2015) proposed a new approach by introducing two techniques, P-DBSCAN and the Markov Chain, and discovered the locations of interest to travelers, their travel patterns, and their daily activities of Hong Kong inbound tourists. Based on data retrieved from Flickr in Austria and all of its regions, Önder (2017) investigated tourist movements during their multi-destination trips using cluster analysis. Travel-related information extracted from geotagged photos can be adopted to customize trip
recommendations via promoting pictured attractions. For example, Lu, Wang, Yang, Pang, and Zhang (2010) proposed a framework to provide a customized trip plan for a tourist, including the popular destinations to visit, the visiting order of destinations, the time arrangement in each destination, and the typical travel path within each destination. By considering user’s preference, present location, spare time and transportation means Kurashima, Iwata, Irie, and Fujimura (2012) proposed a travel route recommendation method to personalized travel plans for tourists.

Despite various studies exploring geotagged photos in tourism research, most of them have been conducted at a city, state, or country level, and few studies have been done at a scenic spot scale, especially at a national park scale in the U.S. This paper investigated the spatiotemporal characteristics of tourist preferences in the most visited national park in the U.S and focused on making contributions to knowledge about tourist behaviors and preferences within the national park system.

3. Data and Methods

3.1 Study Area

The Great Smoky Mountains National Park (GSMNP) is a national park located along the Tennessee–North Carolina border in the southeastern United States. The GSMNP is a treasure of great historical and natural value, offering numerous heritage sites and natural attractions. There are five visitor centers located in the park, including Gatlinburg Welcome Center, Sugarlands Visitor Center, Clingman’s Dome Visitor Center, Cades Cove Visitor Center, and Oconaluftee Visitor Center, in which the Sugarlands Visitor Center and Oconaluftee Visitor Center are the two main visitor centers because they are near the two main entrances to the park: the Gatlinburg entrance in Tennessee and the Cherokee entrance in North Carolina. The U.S. Highway 441 is one of the main roads in the GSMNP and bisects the park, providing automobile access to many trailheads and overlooks. The park includes nearly 850 miles of trails and unpaved roads for hiking, which provides outdoor recreation activities for the benefits of park visitors (Figure 1).
3.2 Data Collection

In this study, Flickr was chosen as the main source of geotagged photos data due to the ease of accessing photos through a public application programming interface (API) and its prevailing popularity among all photo-sharing websites. As mentioned, Flickr’s API was used to retrieve the data (the metadata of geotagged photos) (https://www.flickr.com/services/api/), including textual information (e.g., title, photo tags), geographical information (e.g., longitude, latitude, and name of location), date information (the date when the photos were taken and uploaded), and additional information (e.g., user name, server name, photo ID).

The data of geotagged photos taken in or near Great Smoky Mountains National Park over the past decade from 2007 to 2016 were collected to investigate tourists’ preferences and behaviors. Applying a thorough Flickr database query using keywords such as “smoky mountains”, “great smoky mountains” or “great smoky mountains national park” as well as geolocations, a total of 54,010 photos were collected and used in this study. Given that visitors are able to upload multiple photos at the same location at the same time although some of photos can be anchored in the wrong places in Flickr, we cleaned the data by deleting photos posted with coordinates outside the boundaries of GSMNP and kept the metadata of one photo if there were multiple photos uploaded by the same person at the same time and location. We retrieved a total 7,090 photos from 626 people for the purposes of this study.

3.3 Methods

Geographic Information System (GIS) techniques were applied as the primary analytical method in this study. To investigate the spatial patterns of geotagged photos in the GSMNP, kernel density analysis (KDE) was performed in ArcGIS to identify the hotspots. KDE is widely used in hotspot mapping in tourism research (Ahas, Aasa, Roose, Mark, & Silm, 2008; Beeco & Brown, 2013; Lyon, Cottrell, Siikamäki, & Van Marwijk, 2011) and it creates continuous surfaces of spatial densities with point or area data (Lyon et al., 2011). It divides the space into cells given a user-defined cell size and calculates the density value with a Kernel function (such as Gaussian kernel, Quartic kernel). The area with a high-density value is considered as hotspot. Then, we create a buffer zone along the trail and road, and compute the geotagged photos associated with each trail or road to identify popular trails and roads, based on the assumption that tourists tend to take and share photos at locations they like and enjoy. It is commonly recognized by prior psychological studies, which identify geo-tagging as the same means of self-impression on social media (Gilbert & Barton, 2013; Kim, 2016). In addition, the temporal distribution patterns of geotagged photos in the GSMNP were analyzed and visualized at scales including yearly, monthly, weekly, and hourly distribution patterns. Movements of tourists in the GSMNP were obtained within the ArcGIS from spatiotemporal patterns of geotagged photos. We wrote a script to create multiple lines to link visitors’ locations by the order of photo timestamp. Time sequence movements of tourists were established in order to understand spatiotemporal characteristics of tourists in the GSMNP.

4. Results

4.1 Spatial Characteristics

The results of the kernel density analysis of geotagged photos in GSMNP posted from 2007 to 2016 are presented in Figure 2. As is shown, the light colors to the darker colors indicate the low density of photos to a higher density of photos. Multiple clusters shown in the figure indicate that there were multiple hotspots or attractions in GSMNP such as Clingman’s Dome, Cades Cove Road and Roaring Fork, Newfound Gap, etc. Also, to verify the results, we compare the pictured hotspots results with the top ten popular attractions ranked by users of TripAdvisor, the largest travel website in the world with more than 630 million reviews covering about 7.5 million accommodation, attractions, restaurants and other travel-related businesses (TripAdvisor, 2018). It is found that the results of pictured hotspots were in consistent with popular attractions in the GSMNP ranked by users of TripAdvisor (2017), demonstrating the great potential of implementing geotagged photos data in investigating the tourist preferences for attractions.
Figure 2. The density of photos in the GSMNP

Assuming that the roads and trails of the GSMNP with the most geotagged photos would be popular spots that attracted tourists most, we employed ArcGIS to compute and visualize those geotagged photos within a 100m buffer of roads and trails (Figure 3). In Figure 3 (a), we were able to visualize roads, like Newfound Gap Road North, Clingman’s Dome Road, Cades Cove Loop Road, and Cable Mill Road, which are the four popular roads that consistently see higher visitation and traffic congestion during peak seasons. Figure 3 (b) shows popular trails that attract tourists for hiking, including Sugarland Mountain Trail, Chimney Tops Trail, Appalachian Trail, Noland Creek Trail, and Two-Mile Lead Concession Horse Trail. Park managers and policymakers were able to gain real-time visitors’ behavior and traffic flow information through thousands of geotagged photos posted along the popular roads and trails. The U.S. National Park System may consider implementing more effective strategies to better manage its national park units.
Figure 3. The density of photos along the Roads (a) and Trails (b)
4.2 Temporal Characteristics

The results of the temporal analyses of Flickr geotagged photos retrieved from GSMNP visitors are shown in Figure 4. In Figure 4, yearly, monthly, weekly, and hourly distributions of geotagged photos are presented. More photos were uploaded during the years from 2007 to 2016, which may relate to an increasing number of users of Flickr. The monthly distribution shows a two-peak pattern: one peak is in April, and the other is in October. This trend is relatively consistent with the best visit time of GSMNP due to spring/Easter holidays and fall break/autumn leaf color changes. Photos are mainly uploaded during weekends, which is consistent with visitors’ leisure time with friends and families. For the hourly distributions, a majority of the photos were uploaded between 8:00 am to 18:00 pm. There are also two peaks: one is around 11:00 am to 12:00 noon, and the other around 17:00 pm to 18:00 pm, which is typically consistent with the visitors’ activity patterns.

![Image (a)](image1.png)  
(a) Yearly distribution of geotagged photos.

![Image (b)](image2.png)  
(b) Monthly distribution of geotagged photos.

![Image (c)](image3.png)  
(c) Weekly distribution of geotagged photos.

![Image (d)](image4.png)  
(d) Hourly distribution of geotagged photos.

Figure 4. Analysis of the temporal distribution of Flickr geotagged photos

4.3 Spatiotemporal Characteristics

Geotagged photos presented both location and time information to investigate movements of tourists in national parks; we constructed time sequence movements of tourists in the GSMNP (Fig. 5). Multiple polylines were created to link visitors’ locations along with the order of the timestamps of photos. Regions and paths with many lines passing through the mapped area indicate a high density of movement. The figures showed that there was a high density of movement in the central area of the GSMNP because this area has multiple attractions and are highly connected to scenic spots. Specifically, four main locations were connected with multiple lines, which are the Cades Cove Road, Cataloochee Entrance Road, Sugarlands Visitor Center, and Oconaluftee Visitor Center. It is worth noting that Clingman’s Dome and Newfound Gap were found to be the two most important hubs connecting park tourists from all directions.
5. Conclusion

This study analyzed and visualized the spatial and temporal patterns of Flickr geotagged photos of the Great Smoky Mountains National Park from 2007 to 2016. Several hotspots or attractions in GSMNP were identified and the most popular roads and trails that attracted visitors in GSMNP were detected using the geotagged photos as well. The number of photos uploaded increased annually, and more photos are taken and uploaded by visitors during weekends. The results of this study can be used to identify popular tourism hotspots in national parks, as well as provide practical implications for the Great Smoky Mountains National Park tourism managers in terms of destination development and logistical management. Limitations exist for this study. First, the limited number of users and photos that were uploaded via Flickr geotagged photos only represent a small proportion of the whole population that visited GSMNP. Second, only Flickr data was used in this study, other social media data sources such as Twitter and Instagram may be included to examine the differences among various data sources. Finally, more data sets may be retrieved across different national parks across other states in the U.S. to more comprehensively investigate national park visitors’ behaviors and preferences.

This study is important both theoretically and practically. It contributes to geotagged photos research with focusing on exploring the tourists’ preferences in national park. Also, the methodology used in this study is not limited to Flickr, and can be applied to other social media platforms, like Twitter, Instagram. On a practical level, this study enables tourism planner and managers in national park to implement effective management strategies, such as optimize the allocation of tourist facilities, better understand tourist demand and preference for the sake of sustainable development strategically.

References


Copyrights
Copyright for this article is retained by the author(s), with first publication rights granted to the journal.
This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).