

Co-occurrence Analysis of User-generated Tourism Landscapes in Social Media

Xuan ZHANG

Master/Department of Tourism Administration
Kangwon National University, Chuncheon, ROK

Hee Jeong YUN (Corresponding author)

Professor/Department of Tourism Administration
Kangwon National University, Chuncheon, ROK

Abstract—Although research in the tourism literature that focuses on social media has increased, it seldom addresses landscape, one of the most powerful tourism attractions. Through a content analysis and co-occurrence analysis of posts on microblogs, one of the most popular social media in China, this article links landscape with tourism and social media and explores the online distribution and characteristics of tourism related to landscape. The results of this study show that terms related to cultural landscape in the northern area occur more frequently than they do in the southern area and that, conversely, terms related to natural landscape occur more frequently in the southern area. These results also indicate the differences between the landscape components of the southern and northern areas in China, thereby providing a landscape distribution outline for potential tourists and reminding tourism managers to pay attention to landscapes as tourism attractions.

Index terms -Online tourism landscape, Landscape resources, Cultural landscape, Contents analysis, Microblog

I. INTRODUCTION

The application of social media has widened rapidly to encompass virtually all quarters of daily life. Kaplan and Haenlein (2010: 61) defined social media as “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated content.” This kind of social media includes blogs, collaborative projects, social networking sites, virtual game worlds, and virtual social worlds (Kahar, Yamimi, Bunari, & Habil, 2012). Currently, the most prevalent social media in the world are Microblogs like Twitter, Facebook, and YouTube. Microblog posts can include such formats as plain text, hypertext, images, and digital video and/or audio and can be generated with a mobile device (Karmarkar & Peters, 2013). Because of microblogs’ convenience and mobility and the diversity of content that can be uploaded, the number of microblog users is increasing rapidly. Microblogs are also among the most popular social media in China. Descended in 2006 from blogs, microblogs, one of the most influential social media in China (Li, & Li, 2013), are effective platforms for news releases, advertising, information, and public opinion. National microblog teams have also become prevalent as increasing numbers of government agencies join the microblog trend to convey, collect, and communicate information. Since tourism is an

information-intensive industry, microblogs as social media help disseminate information about tourism to tourists before travel, in the travel sites, and after travel. Therefore, many official tourism sites, which are usually administered by regional tourism bureaus, use microblogging to release regional tourism resources and information. Although some of the literature related to tourism has paid attention to social media, the research in this area remains insufficient. Most of these studies have concentrated on social network services (SNS) like Facebook and Twitter or government websites (Horng & Tsai, 2010; Jankowski, Andrienko, Andrienko, & Kisilevich, 2010; Hsu, 2012; Stepchenkova & Zhan, 2013), but only a few studies have focused on Chinese microblogs. Moreover, most papers related to social media have focused on analyzing the image of tourism destinations (Choi, Lehto, & Morrison, 2007; Stepchenkova & Zhan, 2013), while rarely focusing on the distribution of online tourism landscape in a certain area, although a few studies have examined the co-occurrence frequency of different landscape attributes (Stepchenkova & Zhan, 2013). In addition, although geographic and environmental sciences studies have often explored the categories of landscape, as an industry based on various tourism landscapes, the tourism field has paid little attention to such landscape research. For these reasons this study explores the distribution and frequency of tourism landscapes mentioned in official online tourism microblogs. Then content and co-occurrence analyses are used to examine each of two tourism landscapes, one in northern China and one in southern China. Significant results are provided for diverse stakeholders, tourism planners, tourism managers, and government policy-makers regarding the development of tourism resources that use natural and cultural landscapes.

II. LITERATURE REVIEWS

A. Tourism Landscape

“Landscape” refers to “a setting for human experience and activity as a broad sense [that is] larger than a household but smaller than one of earth’s biogeographical regions” (Riley, 1992, p 13). Therefore, landscape is closely related to place attachment, an affective relationship between people and the landscape that goes beyond cognition (Riley, 1992), and to ordinary life as the cultural meaning of landscape (Meinig, 1979). Landscape is considered a common

resource for the public, so it is susceptible to overuse and has limited carrying capacity (Healy, 1994). Even so, landscape or scenic beauty is an obvious tourism attraction (Inskip, 1991: 94). Knudsen, Metro-Roland, Soper, and Greer (2008) suggested that 'landscape' is the result of interplay between nature and culture and the interplay between natural forces and history, and the term 'landscape' is widely applied in the ecology, environmental science and geography fields. Scholars in these fields generally observe that four factors—physiography, geological formations, vegetation and water resources—determine the characteristics of landscape, in the meantime, regional cultural activities also influence the character of landscape to some extents (Mougiakakou, Tsouchlaraki, Cassios, Nikita, Matsopoulos, & Uzunoglu, 2005). Many studies related to ecology have focused on the classification of landscape, producing various classification or assessment models based on various geographical environment and analysis objects (Strand, 2011; Mougiakakou *et al.*, 2005; Roba & Oba, 2009). However, in the tourism field, classification of landscape is often based on regional resources, such as cultural and natural resources (Deng, King, & Bauer, 2002), and commercial tourism often takes natural and cultural landscapes as the key attraction (Buckley, Ollenburg, & Zhong, 2008). Kent and Elliott (1995) divided landscape into three types: natural landscape, cultural landscape, and mixed landscape (a combination of natural and cultural resources).

According to UNESCO World Heritage Centre (2013), cultural landscape is developed by the interaction between nature and human beings, which shows the evolution of human society under the influence of its natural environment, society, and economic and cultural forces. UNESCO World Heritage Centre divides the cultural landscape into three main categories: garden and parkland, which is created by human beings for aesthetic purposes; organically evolved landscape, including relics or fossils and continuing landscapes that are linked with traditional lifestyles and reflect the form and components of the evolution process; and associative cultural landscape, which may be not belong to the material culture. On the other hand, selecting national parks as an example, Deng *et al.* (2002) defined cultural landscape as consisting of architecture, religion and history, and Kent and Elliott (1995) defined historical residences, colonial villages, farm buildings and so on as the cultural landscape. Natural landscape often involves physical environmental resources, and is related to other natural factors such as climate, topography, wildlife and vegetation, geology and invisible ecological system (Inskip, 1991: 56-58). Many studies have shown that natural landscape is the key attraction of tourism destinations. Brown and Brabyn (2012) confirmed people's preference for natural landscape, while Velarde, Fry, and Tveit (2007) not only showed that people have a preference for natural landscape but also demonstrated the positive health effects of natural landscapes. Most studies have listed natural resources like vegetation, water, mountains, rivers and hills as belonging to

natural landscape (Kent & Elliott, 1995; Brown & Brabyn, 2012). In a similar vein, a classification of landscape can be useful in understanding the tourism landscape and the development of tourism destinations. Therefore, it is useful to add up the landscape and scenery nouns and to examine the associations between them.

B. Social Media

Recently, many studies related to tourism are established in or based on social media (Xiang & Gretzel, 2010; Hsu, 2012; Kisilevich, Ang, & Last, 2012; Boley, Magnini, & Tuten, 2013; Sparks, Perkins, & Buckley, 2013; Stepchenkova & Zhan, 2013). Social media is not only an important tool for travelers who wish to obtain information (Xiang & Gretzel, 2010) but also a new platform for governments, hotels or travel agencies that wish to advertise to or serve the public. The tourism literature related to social media falls primarily into two types: That which concentrates on tourism managers or administrators generally investigates the information and resources collected from government, hotel, travel agency and similar websites (Choi *et al.*, 2007; Hus, 2012). The second type focuses on users of social network websites like Facebook, Flickr, and Twitter in order to understand tourists' behavior or preferences (Jankowski *et al.*, 2010; Ho, Lin, & Chen, 2012). According to these studies, social media reflects the real world, and researchers can draw valid results in terms of SNS. However, a few studies have also investigated the difference between tourists' online posts and tourism managers' posts (Sparks *et al.*, 2013; Stepchenkova & Zhan, 2013). Using a combined approach, Stepchenkova and Zhan (2013) categorized the main attributes of destinations pictured in Flickr's user-generated photos and Peru's official tourism website and conducted a comparative analysis. Although they achieved significant results, they also proposed that, in order to shape the overall image of a certain destination, more research should focus on tourists with different cultural backgrounds or on additional destinations.

III. SCOPE AND METHOD

A. Blog Selection and data collection

Based on the related research, the purpose of this study is to determine the distribution and characteristics of the online tourism landscape. The study uses one of the biggest microblogs in China for this purpose. One of the most popular social media in China is microblogs, this is similar to Twitter. In China in 2012, there were more than 309 million blog users, with those between 10 and 39 years of age constituting nearly 81 percent of the users and 84 percent of the total visit length (Tang, Wu, Huang, & Liu, 2013). Chinese microblogs are dominated by four major web portals: Tencent, Sina, Sohu, and Netease. Tencent blog and Sina microblog are the largest, with 245.3 million and 233.2 million users, respectively (Tang *et al.*, 2013). Among the four microblogs, this study uses the posts of official tourism microblogs in the Sina microblog portal to investigate the contents of the posts and to explore

the distribution of Chinese tourism landscapes. Because Sina microblog website also has authenticated users, like Twitter does, and Tang *et al.* (2013) reported that the number of authenticated users of the Sina microblog portal exceeded 110,000 at the end of 2012. The function and use of these microblog websites are similar to those of Twitter, where each posting allows users to post 140 characters, and users can follow others, leave comments for others, and so on. And to ensure the reliability of data collection, this study used text information posts posted by authenticated tourism users in the Sina microblog website (<http://www.weibo.com>). Authenticated tourism users or accounts in this study refer to accounts created and administered by the regional tourism bureau.

China's natural geography divides the country into four geographic areas: the Qinghai-Tibet area, the Northwest area, the Southern area and the Northern area (Figure 1). Because most provinces in China belong to the Southern and Northern areas and other two areas have totally different landscape, so it's not appropriate of differential analysis between areas, this article explores the landscape distribution in Southern and Northern areas. This range of exploration is immense, so data collection proceeded as follows. First, the study selected ten provinces' authenticated tourism accounts from the Southern area and ten from the Northern area respectively, based on their authenticated tourism accounts' activity levels. The authenticated tourism accounts selected for the Southern area were those of Zhejiang, Shanghai, Hubei, Fujian, Jiangxi, Hunan, Guizhou, Chongqing, Sichuan and Guangxi provinces, and the accounts selected for the Northern area were those of Beijing, Tianjin, Ningxia, Jilin, Liaoning, Hebei, Henan, Shandong, Shanxi, Shaanxi provinces. Because of the large number of posts on these microblogs, this study selected twenty textual posts posted before June 2013 from each province's authenticated tourism account, so the total number of posts collected was 400, half from the Southern provinces and half from the Northern provinces. Each textual post was associated with introducing a scenic spot and contained at least one scenery noun, such as 'river,' 'lake' or 'museum'.

B. Analysis methods

This study employs content analysis and co-occurrence analysis. In the first stage, the content analysis, the researchers classified all postings according to the scenery-related nouns that appear in them. Each textual post contains one or several scenery nouns. Then a frequency analysis obtains an overview of the scenery nouns and the result of 400 textual posts' codings in the southern and northern areas of China. The co-occurrence analysis is used to demonstrate the frequency of co-occurrence and landscape distribution characteristics in southern and northern China. Co-occurrence analysis is used to test the frequency with which two attributes co-occur (van Osselaer & Janiszewski, 2001; Li & Stepchenkova, 2012). As Li and Stepchenkova (2012) explained, if K and L are two independent attributes, then the frequency of co-occurrence

between K and L is a random variable binomially distributed. The meantime z-score is generally the mean of 0 and the standard deviation of 1, and in co-occurrence analysis, the critical z-score is 1.96 when the significance level of two-directional hypothesis testing is 0.05. The calculation of the z-score is the difference in value between the actual and expected co-occurrences of K and L, divided by the standard deviation of K and L's frequency (Stepchenkova & Zhan, 2013). As Stepchenkova & Zhan (2013) indicated that large and positive z-score shows a positive association between two attributes, whereas the large and negative z-score would mean the negative association. Data analysis was conducted using SPSS 17.0 and co-occurrence analysis, was organized into above two parts.



Figure1. Four geographic areas of China

IV. STUDY RESULTS AND DISCUSSION

A. Frequency

As Table 1 shows, 1,091 scenery nouns appeared in the 400 postings: 449 natural nouns and 642 cultural nouns, and 612 nouns in the southern area and 479 nouns in the northern area. The percentage of cultural landscapes in the northern area is higher than that in the southern area, possibly because cultural resources remain comparatively pristine in the northern area. The frequencies between natural nouns and cultural nouns in the southern area are similar to those in the northern area. The frequency analysis revealed 27 scenery nouns are deducted: mountain (151), river (70), plant (48), forest (39), lake (33), canyon (27), sea (20), karst cave (18), waterfall (18), island (15), and hot spring (10), all belong to the natural landscape, while culture (187), historical architecture (142), religion (46), custom (38), park (32), ethnic minority (29), revolutionary site (28), temple (28), ancient city (28), rural area (23), museum (16), activity (14), art (13), food

(9), rafting (5), and shopping (4) all belong to the cultural landscape. Among the cultural nouns, the terms related to culture contain both historical concepts and contemporary human life. As total frequencies show, culture (187), mountain (157) and historical architecture (151) rank as the top three most frequently used terms. Shopping (4), rafting (5), food (9) rank in the bottom three, with shopping appearing only four times, perhaps because although shopping generally exists in most scenic spots, but only a few scenic spots regard shopping as the main attraction. Based on these results, the terms of cultural landscape vary more than do those of natural landscape in Chinese microblogs because the frequencies of nouns related to cultural landscape are distributed widely and perhaps also because these results are related to geographical resources, regional development, and identities. In addition, some scenery nouns, such as activity, custom, rafting, religion, and shopping, are related to human activities or beliefs based on physical resources. Therefore, invisible factors are also important in the tourism landscape because tourism is an activity-based field, and potential tourists see these invisible resources as one of the resources of the physical landscape resources or do not divide landscape into visual and invisible landscape.

Table 1 also shows the frequency which these terms appear in the posts from the southern area and the northern area. Like the result obtained from analyzing all 400 posts, the southern area's top three tourism landscapes are mountain (97), culture (79), and historical architecture (58), while the northern area's top three tourism landscapes are culture (108), historical architecture (84), and mountain (54). Moreover, culture and historical architecture appear more often in posts from the Northern area than in those from the Southern area. This difference may be due to most ancient capitals in Chinese history being located in the Northern area. What's more, although 'rafting' appears only five times in the Southern area's posts, but no posts from the Northern area mention it, perhaps because, the water resources are more abundant in the Southern area. "Ethnic minority" and "ancient city" also appear more often in the southern area's posts than in those of the northern area, perhaps because the ethnic minority is distributed more widely in the southern area, particularly in the provinces in the southwest of China. With respect to the ancient city, the co-occurrence analysis provides valid reasons for why it appears more often in the southern area's posts than in those of the northern area. Furthermore, since the southern area is rich in mountainous topography, the natural nouns like canyon and karst cave also appear more frequently in the southern area than the northern area.

Table1. Result of frequency analysis of 400 textual posts

| Scenery nouns (classification of landscape) | Frequencies (%) | | |
|---|------------------|----------------|----------------|
| | Total | Southern area | Northern area |
| Activity (C) | 14 (3.5) | 6 (3.0) | 8 (4.0) |
| Ancient city (C) | 28 (7.0) | 23 (11.5) | 5 (2.5) |
| Art (C) | 13 (3.3) | 8 (4.0) | 5 (2.5) |
| Canyon (N) | 27 (6.8%) | 21 (10.5) | 6 (3.0) |
| Culture (C) | 187 (46.8) | 79 (39.5) | 108 (54.0) |
| Custom (C) | 38 (9.5) | 21 (10.5) | 17 (8.5) |
| Food (C) | 9 (2.3) | 8 (4.0) | 1 (0.5) |
| Forest (N) | 39 (9.8) | 22 (11.0) | 17 (8.5) |
| Historical architecture (C) | 142 (35.5) | 58 (29.0) | 84 (42.0) |
| Hot spring (N) | 10 (2.5) | 6 (3.0) | 4 (2.0) |
| Island (N) | 15 (3.8) | 11 (5.5) | 4 (2.0) |
| Karst cave (N) | 18 (4.5) | 15 (7.5) | 3 (1.5) |
| Lake (N) | 33 (8.3) | 20 (10.0) | 13 (6.5) |
| Ethnic minority (C) | 29 (7.3) | 22 (11.0) | 7 (3.5) |
| Mountain (N) | 151 (37.8) | 97 (48.5) | 54 (27.0) |
| Museum (C) | 16 (4.0) | 6 (3.0) | 10 (5.0) |
| Park (C) | 32 (8.0) | 16 (8.0) | 16 (8.0) |
| Plant (N) | 48 (12.0) | 22 (11.0) | 26 (13.0) |
| Rafting (C) | 5 (1.3) | 5 (2.5) | 0 |
| Religion (C) | 46 (11.5) | 24 (12.0) | 22 (11.0) |
| Revolutionary site (C) | 28 (7.0) | 17 (8.5) | 11 (5.5) |
| River (N) | 70 (17.5) | 54 (27.0) | 16 (8.0) |
| Rural area (C) | 23 (5.8) | 15 (7.5) | 8 (4.0) |
| Sea (N) | 20 (5.0) | 9 (4.5) | 11 (5.5) |
| Shopping (C) | 4 (1.0) | 3 (1.5) | 1 (0.5) |
| Temple (C) | 28 (7.0) | 11 (5.5) | 17 (8.5) |
| Waterfall (N) | 18 (4.5) | 13 (6.5) | 5 (2.5) |
| Total frequency | 1,091 (100.0) | 612 (100.0) | 479 (100.0) |
| Frequency of natural landscape | 449 (41.2) | 290 (47.4) | 159 (33.2) |
| Frequency of cultural landscape | 642 (58.8) | 322 (52.6) | 320 (66.8) |

(N): Natural landscape, (C): Cultural landscape

B. Co-occurrences Analysis

Figure 2 shows the result analyzing the scenery nouns from the 400 textual posts of the southern and northern areas using co-occurrence analysis. Nouns that belong to the cultural landscape are represented by dashed bubbles, while natural landscape nouns are represented by solid bubbles. The three most frequent nouns, culture (187), mountain (151), and historical architecture (142), are represented in the middle and at the bottom of the map. For the purpose of keeping the co-occurrence map concise, only actual co-occurrence values and z-scores, not the expected co-occurrence values of each two nouns, are shown. Values in parentheses are z-scores, so the more positive and larger the value, the more positive the link

between the two nouns (Li & Stepchenkova, 2012). In addition, because the critical z-score is 1.96, only values higher than +1.96 are shown. For example, the number 26 between “religion” and “temple” indicates that 26 textual postings contain both words. Since the expected co-occurrence of these two nouns is 3.22 (not shown), the z-score is 12.75, indicating the strength of statistical association between these two attributes.

With a value of 26 actual co-occurrences, “temple” and “religion” have the largest z-score value (12.75) on the map, indicating that these two attributes are positively associated. The high-frequency noun “mountain” (151) is positively associated with seven other nouns (river, canyon, waterfall, forest, lake, park, and religion), indicating that the landscape resources in mountainous areas are diverse. That the expected co-occurrence number between “mountain” and “forest” is 14.72, the actual number is 35, and the relationship of these two attributes has the highest z-score (5.39) among seven nouns indicates that “forest” has the most positive link to “mountain.” “Forest” also has the large number of significant links to other attributes, while “park” is the most commonly associated attribute. Although “park” is classified as a cultural landscape in this paper, “forest” and “park” have a positive linkage here. The relationship between “culture” and “historical architecture” has the highest actual co-occurrence number (117), so the expected number between them is 66.38, and the z-score is 6.8. Therefore, there is a highly positive association between them, indicating that “culture” and “historical architecture” are entangled with each other in scenic spots. Although the frequencies of “culture” and “historical architecture” ranked first and second, their low z-score links show that these two nouns are rarely positively associated with other nouns, so most scenic spots that have both culture and historical architecture lack diversification. A possible reason for the strong positive link between “ethnic minority” and “custom” (z-score: 10.43) is that the most attractive feature of ethnic minority areas is ethnic customs. The co-occurrence value between “river” and “canyon” is 15, with a 4.76 z-score, indicating that rivers frequently appear near canyons. The other notable attribute pair is “museum” and “art”; there is a positive association between these two attributes but no valid link with other attributes. “Rafting” appears with low frequency in the sample, but it has positive links with other nouns, such as “river,” “canyon,” and “hot spring.” “Island” and “sea” are highly associated, as expected, with a 7.22 z-score. “Shopping” (4) has no positive association with other nouns.

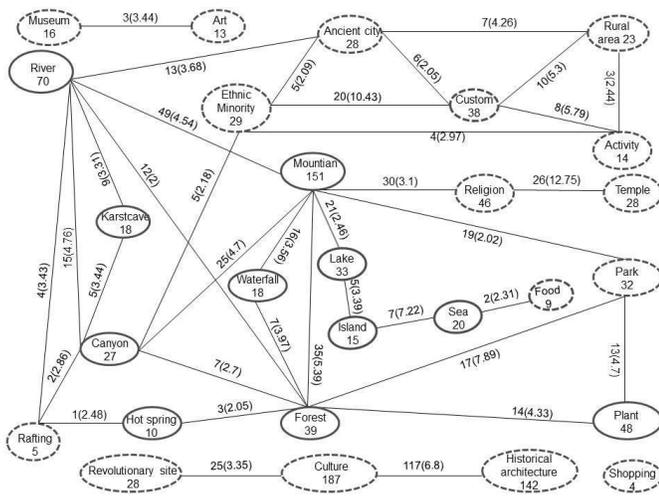
Figures 3 and 4 illustrate the co-occurrence associations in the southern and northern areas, respectively. The co-occurrence maps show that the noun “mountain” has the highest number of valid links in both the northern area and the southern area. In the southern area, mountain (97) has valid links with canyon (21), river (54), waterfall (13), forest (22), and lake (20). Among them, “mountain” and “forest” had the most positive z-score (4.89). In the northern area, although

“forest” and “mountain” also had the most positive associations among six significant links, “temple” and “religion” and the relationship between “religion” and “mountain” had a strong positive association (z-score: 4.19). In the northern area “religion” linked with other four attributes (mountain, temple, forest, and shopping), while in the southern area religion had a significant link only with “temple,” indicating that religion-related scenic spots in the northern area are more diversified. The co-occurrence of “culture” and “historical architecture” is similar to the overall result of the southern and northern areas; although these two attributes appeared often both areas, they did not have valid links with other attributes. Nevertheless, in the southern area, “historical architecture” connects to “ancient city” with 12 co-occurrences and a z-score of 2.1, perhaps because there are more ancient cities in the southern area than in the northern area and the main attraction of ancient cities is their historical streets or residential buildings. “Ancient city” also presented a positive association with “river” in the southern area, possibly because the southern area is rich in water resources and many ancient cities were built near water/streams. “Revolutionary site” has only one valid link (with “culture”) in both the southern and the northern areas, possibly because revolutionary sites are nearly always surrounded by a serious atmosphere aimed at showing the culture and old sites related to revolutionary history; hence, these areas have a low association with other tourist sites. “Rafting,” “hot spring,” “art,” and “lake” have no significant links with other attributes in the northern area map. In particular, “rafting” had 0 frequency in the northern area. However, these terms all had positive links in the southern area. For example, the z-scores between “rafting” and “river,” “canyon,” and “hot spring” are all higher than 2.0 in the southern area. In addition, “lake” connects to “island” and “mountain” in the southern area; the z-score for “lake” and “island” was nearly 4, while it had no significant link in the northern area, perhaps because the longer coastline in the southern area denotes more islands and sea landscape. What’s more, “shopping” and “food” had positive links in the southern area, whereas in the northern area “shopping” is entangled with “temple” and “religion.” “Custom” and “ethnic minority” has the largest z-score (7.74) in the southern area, while in the northern area “temple” and “religion” had the most positive association (9.65). All of these indicate that there are differences in landscape composition between the southern and northern areas.

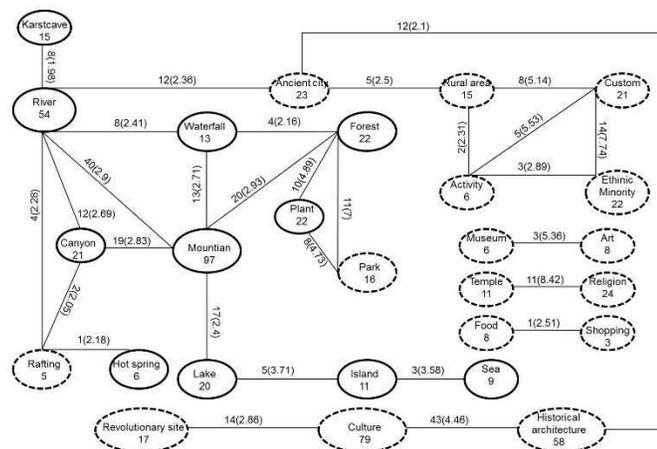
V. CONCLUSION

For the purpose of exploring the most frequent scenery nouns posted by Chinese authenticated tourism accounts to determine the landscape distribution in Southern and Northern area in China, this study selected authenticated tourism accounts in the Sina microblog portal created and administered by regional tourism bureau as a study object and collected their textual posts associated with introducing scenic spots as the analysis data. (Each post contained at least one scenery noun.) The results show that the scenery nouns posted most frequently are ‘culture’, ‘mountain’ and ‘historical architecture’ which indicate that the Southern and Northern area of China is abundant in terms of these three tourism resources. Frequency analysis shows that cultural landscape nouns appear more often in the northern area (66.8%) in Chinese microblogs, while the numbers of cultural landscape nouns and natural landscape nouns are almost balanced in the southern area (52.6% and 47.4%, respectively). Reasons for this result may be that most Chinese historical capitals are located in the northern area, so the frequency of “culture” and “historical architecture” in the northern area’s microblogs significantly outnumbers the frequency in the southern area’s microblogs. “Ethnic minority,” “mountain,” and “river” appear more frequently in the southern area because of differences in minorities’ and resources’ geographical distribution. Then the results of the southern and northern area’s comprehensive co-occurrence analysis demonstrate that, among all of the co-occurrence relationships between all the scenic posts, ‘mountain’ has the most active co-occurrence relationship because it is significantly associated with seven other attributes, while the highest z-score, which between ‘religion’ and ‘temple,’ indicates that these two attributes have the most positive association. Separate analyses of the southern and northern areas indicate that landscape distribution and composition in the southern area differ from those in the northern area, as the co-occurrence relationships between the scenery nouns differ significantly. Moreover, to some extent, a higher number of nouns with no associations in the northern area (rafting, hot spring, art, and lake) indicate that landscape in the southern area is more complex. In addition, combining the three maps of co-occurrences shows that cultural landscape nouns are significantly associated with other cultural landscape nouns in each map, and natural landscape nouns co-occur with other natural landscape nouns. Only a few natural landscape nouns show positive links with cultural landscape nouns (park, plant, and forest).

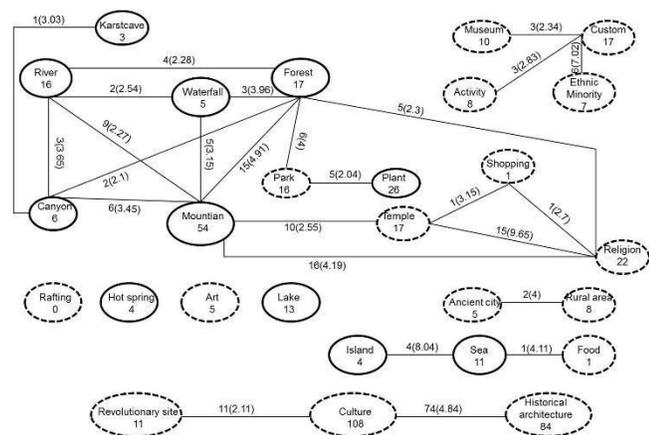
In summary, findings from the co-occurrence analysis show the co-occurrence probability of each of the features of cultural and natural landscapes in the southern and northern areas, provided an outline of landscape distribution for readers who wish to understand the general landscape of the Southern and Northern areas in China and may provide necessary information for travelers when they making tourism plans, and what’s more, may provide valuable information for tourism planning, resource protecting and statistics of tourism



Cultural landscape: solid line, Natural landscape: dotted line Figure2. Co-occurrences map (comprehensive data of the Southern and Northern areas)



Cultural landscape: solid line, Natural landscape: dotted line Figure3. Co-occurrences map (Southern area)



Cultural landscape: solid line, Natural landscape: dotted line Figure4. Co-occurrences map (Northern area)

resource. Nevertheless, researchers should collect more textual posts and do additional co-occurrence analyses of the Southern area and the Northern area in order to develop a more specific map of the landscape distribution in these areas. Future research should conduct further studies of the landscape distribution in whole China in order to get more comprehensive tourism resource information of China for tourism planning, resource protecting and statistics of tourism resource, and should also do further researches of the landscape distribution in areas other than China in order to give a tourism guide information for travelers.

Funding

This study was supported by a research grant from the National Research Foundation of the Republic of Korea (Grant number: NRF-2011-0014120).

REFERENCES

- [1]. Boley, B. B., Magnini, V. P., & Tuten, T. L. (2013). Social media picture posting and souvenir purchasing behavior: Some initial findings. *Tourism Management*, 37, 27-30.
- [2]. Brown, G. & Brabyn, L. (2012a). The extrapolation of social landscape values to a national level in New Zealand using landscape character classification. *Applied Geography*, 35, 84-94.
- [3]. Brown, G. & Brabyn, L. (2012b). An analysis of the relationships between multiple values and physical landscapes at a regional scale using public participation GIS and landscape character classification. *Landscape and Urban Planning*, 107, 317-331.
- [4]. Buckley, R., Ollenburg, C., & Zhong, L. (2008). Cultural landscape in Mongolian tourism. *Annals of Tourism Research*, 35(1), 47-61.
- [5]. Choi, S., Lehto, X. Y., & Morrison, A. M. (2007). Destination image representation on the web: Content analysis of Macau travel related websites. *Tourism Management*, 28, 118-129.
- [6]. Deng, J., King, B., & Bauer, T. (2002). Evaluating natural attractions for tourism. *Annals of Tourism Research*, 29(2), 422-438.
- [7]. Healy, R. G. (1994). The "common pool" problem in tourism landscape, *Annals of Tourism Research*, 21(3), 596-611.
- [8]. Ho, C., Lin, M.-H., & Chen, H.-M. (2012). Web users' behavioural patterns of tourism information search: From online to offline. *Tourism Management*, 33, 1468-1482.
- [9]. Horng, J.-S., & Tsai, C.-T. (2010). Government websites for promoting East Asian Culinary tourism: A cross-national analysis. *Tourism Management*, 31, 74-85.
- [10]. Hsu, Y.-L. (2012). Facebook as international eMarketing strategy of Taiwan hotels. *International Journal of Hospitality Management*, 31, 972-980.
- [11]. Inskeep, E., (1991). *Tourism Planning: An Integrated and Sustainable Development Approach*, John Wiley & Sons, INC.
- [12]. Jankowski, P., Andrienko, N., Andrienko, G., & Kisilevich, S. (2010). Discovering landmark preferences and movement patterns from photo postings. *Transactions in GIS*, 14(6), 833-852.
- [13]. Kahar, R., Yamimi, F., Bunari, G., & Habil, H. (2012). Trusting the social media in small business. *Procedia - Social and Behavioral Sciences*, 66, 564-570.
- [14]. Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53, 59-68.
- [15]. Karmarkar, A., & Peters, R. (2013). *Context-enriched Microblog Posting*, US Patents (US 8489132 B2), <http://www.google.com/patents/US8489132>.
- [16]. Kent, R. L., & Elliott, C. L. (1995). Scenic routes linking and protecting natural and cultural landscape features: A greenway skeleton. *Landscape and Urban Planning*, 33, 341-355.
- [17]. Kisilevich, S., Ang, C. S., & Last, M. (2012). Large-scale analysis of self-disclosure patterns among online social networks users: A Russian context. *Knowl Inf Syst*, 32, 609-628.
- [18]. Knudsen, D. C., Metro-Roland, M. M., Soper, A. K., & Greer, C. E. (2008). *Landscape, Tourism and Meaning*. Aldershot, UK: Ashgate Publishing.
- [19]. Li, X., & Stepchenkova, S. (2012). Chinese outbound tourists' destination image of America: Part I. *Journal of Travel Research*, 51(3), 250-266.
- [20]. Li, Y-M, & Li, T-Y. (2013). Deriving market intelligence from microblogs. *Decision Support Systems*, 55, 206-217.
- [21]. Meinig, D. W. (1979). *The Interpretation of Ordinary Landscape*. NY: Oxford University Press: 2-4.
- [22]. Mougiakakou, S. G., Tsochlaraki, A. L., Cassios, C., Nikita, K. S., Matsopoulos, G. K., & Uzunoglu, N. K. (2005). SCAPEVIEWER: Preliminary results of a landscape perception classification system based on neural network technology. *Ecological Engineering*, 24, 5-15.
- [23]. Riley, R. B., (1992) Attachment to the ordinary landscape, Book title: Place Attachment (Editors: Altman, I & Low, S. M.), Springer US.
- [24]. Roba, H. G., & Oba, G. (2009). Community participatory landscape classification and biodiversity assessment and monitoring of grazing lands in northern Kenya. *Journal of Environmental Management*, 90, 673-682.
- [25]. Sparks, B. A., Perkins, H. E., & Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tourism Management*, 39, 1-9.
- [26]. Stepchenkova, S., & Zhan, F. (2013). Visual destination images of Peru: Comparative content analysis of DMO and user-generated photography. *Tourism Management*, 36, 590-601.
- [27]. Strand, G.-H. (2011). Uncertainty in classification and delineation of landscapes: A probabilistic approach to landscape modeling. *Environmental Modelling & Software*, 26, 1150-1157.
- [28]. Tang, X. J., Wu, X. X., Huang, C. X., & Liu, R. S. (2013). *Annual report on development of new media in China (2013)*. Social Sciences Academic Press (China).
- [29]. UNESCO World Heritage Centre. (2013). Operational Guidelines for the Implementation of the World Heritage Convention. Retrieved from <http://whc.unesco.org/archive/opguide08-en.pdf#annex3>

- [30]. van Osselaer, S. M. J., & Janiszewski, C. (2001). The ways of learning brand associations. *Journal of Consumer Research*, 28, 202-23.
- [31]. Velarde, M. D., Fry, G., & Tveit, M. (2007). Health effects of viewing landscapes: Landscape types in environmental psychology. *Urban Forestry & Urban Greening*, 6, 199-212.
- [32]. Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, 31, 179-188.



Xuan ZHANG received the Master's degree in Department of Tourism Administration of Kangwon National University, Chuncheon, Republic of Korea, in 2014.



Hee Jeong YUN (Corresponding author) is an assistant professor in the department of tourism administration at Kangwon National University, Republic of Korea and has a first degree in Tourism Management and a Master's degree and a doctorate in Landscape Architecture from Seoul National University.