



Sentiment Analysis of Cultural Heritage Landscape Elements Using Big Data of Online Comments: A Case Study of the Humble Administrator's Garden in China

Qianda Zhuang^{1,2} and Shuzhen Chen³(✉)

¹ Faculty of Design and Architecture, Universiti Putra Malaysia, UPM, 43400 Serdang, Selangor, Malaysia

gs56070@student.upm.edu.my

² College of Agriculture and Forestry, Linyi University, Linyi 276000, China

³ School of Foreign Languages, Linyi University, Linyi 276000, China

chenshuzhen@lyu.edu.cn

Abstract. The big data of online comments has been widely used in fields, however, few research on sentiment of cultural heritage conducted. This study explores the sentiment of tourists about the Humble Administrator's Garden using big data of online comments in Ctrip and Qiongyou websites. Word frequency analysis was used to extract the landscape elements and cluster and sentiment analysis were conducted for landscape elements sets. Sentiment analysis was calculated by model of $l(w) = n(w) a(w) s(w) m(w)$. The study shows eight sets of landscape elements interested by tourists exist, in which humanities elements and water related are most considered. The positive sentiment of these two sets is more than 88%, while visiting service is only 69% and stone is 70%. The study provides a supporting attempt of online comment data in the field of cultural heritage, which contributes to the improvement of management and conservation towards the cultural heritage landscape.

Keywords: Historical landscape · landscape perception · big data · cultural heritage

1 Introduction

Cultural heritage is the wisdom of our predecessors and an important symbol of the city. The protection of cultural heritage will help to protect the historical and cultural memory of the city and enhance the characteristics of the city. Only with good protection can we carry the cultural genes and revitalize history. At present, the research on cultural heritage protection focuses on protection mode, inheritance mode, legislative protection, value research, etc. For example, He (2014) studied the factors affecting the satisfaction of cultural heritage scenic spots through grounded theory [1], and Zhang (2021) explored the value of Suzhou Classical Garden heritage through discourse analysis [7]. Most of

the existing studies are based on traditional research methods, and lack of understanding and protection of cultural heritage landscape from the perspective of tourists.

As an important cultural heritage, classical gardens have become an important scenic spot to develop and popularize garden cultural knowledge to the public. Modern information technology and network big data are gradually introduced into the research of classical gardens. The integration of online platform data can deeply tap visitors' multi-dimensional understanding of scenic spots, and provide a quantitative analysis supplement to the traditional qualitative analysis from the perspective of users. The emotional analysis of the online comment text of classical gardens will help to improve the understanding of tourists' attitudes and views on the cultural heritage of classical gardens. It is helpful to explore the certification degree of tourists for classical gardens, so as to provide scientific decision-making for the protection of cultural heritage.

As one of the four famous gardens in Suzhou, the Humble Administrator's Garden has a large area and many tourists. Therefore, this study takes the Humble Administrator's Garden as an example. By exploring the tourists' cognition of the elements of the classical garden landscape of the Humble Administrator's Garden and the emotional analysis of their respective element sets, the main research questions are as follows,

- 1 What do tourists pay most attention to the landscape elements of the Humble Administrator's Garden?
- 2 what kinds of landscape elements can be divided into? How much attention is paid to it?
- 3 what is the emotional attitude of tourists towards each element set of the Humble Administrator's Garden?

2 Research Methodology

2.1 Study Area

The Humble Administrator's Garden is one of the most famous gardens in China. It is the largest and most famous in Suzhou city and includes almost all the characteristics of traditional gardens. It has been inscribed as a World Heritage Site by UNESCO in 1997. The Humble Administrator's Garden is celebrated by its exquisite layout of pools, rockeries, islets, pavilions and etc. It is the representative and example of classical gardens, which originally built in 1509 during Ming Dynasty and developed to present three characteristic parts, the Eastern Garden, Western Garden and Central Garden, covering about 52000 square meters.

2.2 Data Collection

Tourism websites related to Humble Administrator's Garden include Meituan, Qunar, Qiongyou, Hornet's nest, Catwalk eagle, Ctrip, Tongcheng etc. After one-by-one checking, it is found that the tourist comment data of Ctrip and Qiongyou are relatively rich, complete and reliable. Therefore, we choose these two platforms as the data source. The data collection period is from January 1 to December 31, 2021. The Octopus software with convenient, quick and accurate way was used.

Table 1. The number of comments in each selected platform

Source	Original Number of comments	Number of after proceeding
Ctrip	3001	2788
Qiongyou	475	443
total	3476	3231

The collected data needs to be processed, mainly including removing duplicate, deleting comments with no actual meaning and deleting comments with simple expressions. The amount of data obtained and processed are shown in Table 1.

2.3 Data Analysis

According to previous studies, natural language processing, text analysis [2, 5] were used in this study.

- (1) Word frequency analysis. Through word frequency analysis of all data, we can obtain the landscape theme elements that tourists pay more attention to.
- (2) Cluster analysis. Cluster the theme words of many landscape elements, so as to sum up the landscape element set.
- (3) Sentiment analysis. After the emotional evaluation of the elements in different landscape elements, all kinds of emotional analysis can be obtained, so as to find the emotional tendency of tourists for each cluster [4]. The sentiment analysis used the model

$$l(w) = n(w)a(w)s(w)m(w) \quad (1)$$

in (1), $l(w)$ represents the sentiment value of emotional tendency, and $n(w)$ represents the weight value of negative words. $s(w)$ represents the emotional value of the emotional word, and $a(w)$ represents the sum of the weight values of all degree adverbs before the emotional word. $m(w)$ indicates the relative position between negative words and adverbs of degree before emotional words [4].

- (4) Comparative analysis. The obtained emotional results of each element set are compared with each other horizontally, so as to find the emotional differences of tourists between each element set.

3 Results and Discussions

3.1 Analysis of Landscape Elements Concerned by Tourists

Through word frequency analysis and part of speech tagging, we can get the main landscape elements concerned by tourists in the Humble Administrator's Garden. In this study, the theme words with word frequency greater than 10 are selected as the research sample, and a total of 143 theme words are obtained. More intuitive display, visual



Fig. 1. The word cloud map of landscape elements interested by tourists in Humble Administrator's Garden (Drawn by the author)

Table 2. The top 20 landscape elements interested by tourists and its frequency in Humble Administrator's Garden

Element	Frequency	Part of speech	Element	Frequency	Part of speech
Museum	364	noun	Core	129	noun
Visiting guide	357	noun	Mountains and waters landscape	128	noun
Scenic spot	344	noun	Pavilions, terraces and open halls	115	noun
Guide	310	noun	Lotus	103	noun
Ticket	308	noun	Rockery	82	noun
Architecture	295	noun	Layout	78	noun
Garden	237	noun	Ornamental flowers and trees	75	noun
Design	211	noun	Small bridge	74	noun
History	175	noun	Story	68	noun
Culture	155	noun	Service	68	noun

analysis of word cloud, Fig. 1 The word cloud map of landscape elements interested by tourists in Humble Administrator's Garden can be obtained. The top 20 are sorted according to word frequency, and Table 2 can be obtained.

According to the word cloud and Table 2, the top elements that are highly concerned by tourists including museum, tour guide, ticket, architecture, garden, design, history, etc. The first one is "Museum", which refers to China's first garden theme museum located in the residential area on the west side of the Humble Administrator's Garden. It shows in detail the gardening knowledge of folding mountains, managing water, planting flowers and trees, architectural construction etc. Because the location is close with popularized garden art and free tickets, it also welcomed by tourists. In addition, repeated descriptions are inevitable in other elements, such as "Courtyard" and "Yard",

Table 3. The classification of landscape elements in Humble Administrator’s Garden (part)

Landscape elements set	Number	Elements
Architecture	1241	Museum, Architecture, Pavilions terraces and open halls, Pavilion, Courtyard, Folk houses, Yard, Residential district, Window, Pavilions and pavilions, Waterside Pavilions
Stone	118	Rockery, Taihu lake stone, Mountain stone, Stone
Flora and fauna	902	Gardens, Lotus, Ornamental flowers and trees, Luxuriant, Bonsai, Flowers and plants, Tree, Ginkgo, Bamboo, Mandarin duck, Squirrel
Water	368	Mountain and waters landscape, Small bridge, Pond, Inverted reflection in water, Pool
Spatial Structure	1556	Entrance, Scenic spot, Design, Core, Nature, Open, Land occupation, Walk, Pattern, Position, Route, Space, Well-proportioned, Scale, Four sides, Central section, West, East, Scattered, The Long Corridor, Export, Utilize, Gate, Surround, Arrangement, Corridor,
Humanities elements	1083	History, Culture, Park, Layout, Story, Cheerful, Art, Tradition, The ancients, Allusion, Lasting appeal, Artistic conception, Men of letters, Knowledge, Humanity, Poetic charm, Pingtan
Visiting Service	1108	Visiting guide, Guide, Ticket, Service, Commentator, Commentary, Ticket Office
Season and Meteorology	194	Light and shadow, Spring, Air, Misty rain, Sunshine, Summer, Light rain, Late autumn, Winter, Path, Spring Festival, A sunny day

“Pavilions” and “Pavilion”. Therefore, clustering all elements is necessary to obtain the most attention of tourists.

3.2 Cluster Analysis of Landscape Elements

In order to explore the emotion of tourists for specific kind of landscape elements, we cluster them. According to the research of Peng (1986) [6] and Liu (1979) [3] the main elements of Chinese classical gardens can be divided into six categories: architecture, rocks, plants, water body, spatial structure and humanistic elements. In this study, “explanation”, “tour guide” and “ticket” belong to the category of scenic spot service, and “spring”, “drizzle” and other factors belong to season and Astronomy and meteorology. Therefore, visiting service, season and meteorology are added. After manually matching and clustering each element subject word with the above classification, the landscape element clusters are shown in Table 3 can be obtained.

By clustering and summarizing the above landscape elements, we can get the number of relevant comment data related to the above eight types of landscape element sets,

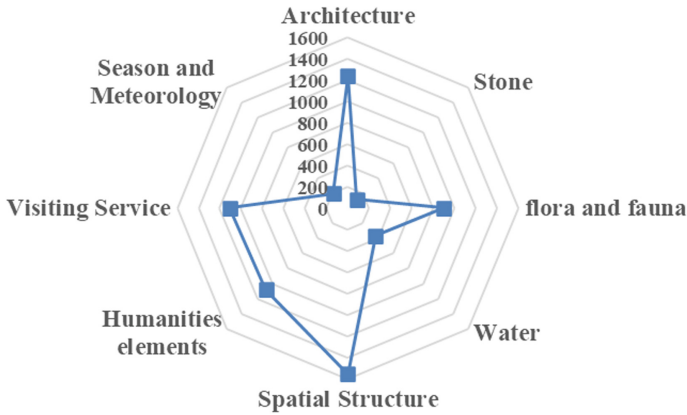


Fig. 2. The degree of each landscape element set that interested by tourists

which can be used as a reference for tourists' attention to each element set. The results are shown in Fig. 2.

According to Fig. 2, tourists pay the highest attention to the spatial structure of the Humble Administrator's Garden, followed by architecture, tourism services and cultural elements. The above four types of attention have exceeded 1000 times. The lowest attention was paid to rocks, only 118 times.

In the spatial structure, the words "entrance", "scenic spot" and "design" have the highest attention, which shows that tourists pay high attention to the landscape setting at the entrance. There are many scenic spots in the park, which are also praised by tourists. The elements of architectural landscape are concentrated, and the three "museums, buildings and Pavilions" pay the most attention. In addition to the Garden Museum, the overall buildings have attracted the attention of tourists, among which the garden buildings such as pavilions are particularly attractive.

In the tour service, "explanation", "tour guide" and "ticket" are the three elements that tourists pay most attention to. Both the explanation and the tour guide show that tourists agree with the explanation of the tour guide during the tour, and can obtain an in-depth understanding of the scenic spot through the tour guide. The concern about "tickets" shows that the ticket price is too high, which has become an aspect of controversy.

3.3 Sentiment Analysis of Each Landscape Element

Each specific element is retrieved from the original text and classified separately, and then a single comment in each category is evaluated by artificial emotion. The description and evaluation of elements are positive, positive and amazing are marked as positive, those described as negative, negative and regretful are summarized as negative evaluation, and those simply described without emotional tendency are marked as neutral. After evaluating the comments of all categories one by one, Table 4 can be obtained. Through the horizontal comparison and analysis, the Fig. 3 can be obtained.

Table 4. Sentiment analysis percent of each element set of Humble Administrator’s Garden

	Positive	Neutral	Negative
Architecture	82	14	4
Stone	70	23	7
flora and fauna	80	16	4
Water	88	9	3
Spatial Structure	82	14	4
Humanities elements	90	5	5
Visiting Service	69	12	19
Season and Meteorology	81	16	3

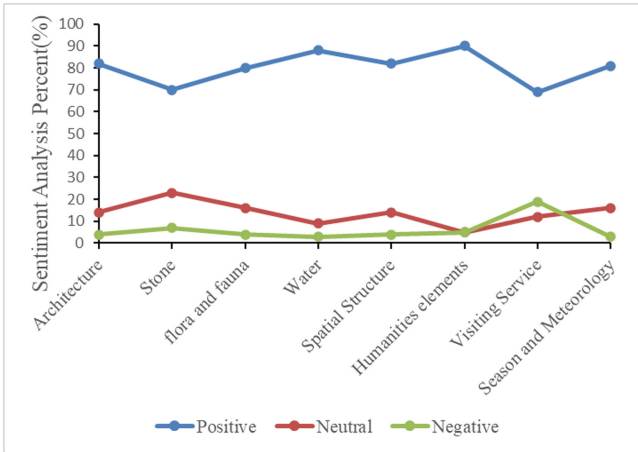


Fig. 3. The horizontal comparison of sentiment analysis of each element set

On the whole, the tourists’ emotional evaluation of the eight landscape elements of the Humble Administrator’s Garden is mainly positive, which reflects the tourists’ agreement of the Humble Administrator’s Garden. However, there are obvious differences exist. Though the positive evaluation reached 90%, 5% said they could not understand the meaning. For example, “although Kunqu Opera and Pingtan are good, we can’t understand the content instead its rhythm”, which requires additional elaborate description, letting tourists have a deeper understanding of the local cultural characteristics.

Compared with other types, the positive emotion of mountains and stones is only 70%, while the neutral ones account for 23%, indicating that some tourists cannot deeply feel the charm of the rockeries and stones in the Humble Administrator’s Garden.

The positive evaluation of animals and plants, buildings and seasonal meteorology is all about 80%. The positive evaluation of tour services is the lowest, only 69%, and the negative evaluation accounts for 19%. This is mainly because the high price of tickets.

Therefore, ticket price should be considered to achieve more tourists' satisfaction. The positive emotion of water body accounts for 88%, indicating that the construction of water body is recognized by tourists. As the most flexible landscape element, it can make people experience the fun of nature.

4 Conclusions

Taking the Humble Administrator's Garden as an example, this paper uses the big data of online comment text as the research data, and obtains a total of 3231 valid data from Ctrip and Qiongyou. Through word frequency analysis, we can get the main themes that tourists pay more attention to. By further clustering these themes according to eight categories: architecture, mountains and rocks, plants, water body, spatial structure, cultural elements, tour services, seasons and meteorology, we can get the degree of tourists' attention to these sets. The study found that tourists pay more attention to the spatial structure, architecture, cultural elements, tourism services and seasonal and weather, and their attention frequency is more than 900.

The individual emotional evaluation of the landscape element data in these clusters can obtain the emotional evaluation of tourists for each element set. On the whole, tourists' emotion towards each element set is positive. However, the differences among sets are also obvious. The positive evaluation of human elements reaches 90%, while the positive emotion of mountains and rocks is only 70% compared with other categories. The positive emotion of water accounts for 88%, and the positive evaluation of animals and plants, buildings and seasonal meteorology is almost the same, all about 80%. This provides a direction for the improvement of the scenic spot in the next step.

This study verifies the availability and effectiveness of Internet tourist comment big data for sentiment analysis of cultural heritage. The content analysis of these data is helpful to obtain the emotional evaluation on the perspective of users, which can provide suggestions for landscape improvement. Moreover, the utilization of such big data demonstrates the trend of the effective integration of traditional gardens and modern information technology. However, some limitations exist. For instance, the sample size of reviews is not big enough and technical barriers exist for data acquisition.

References

1. He Q. F (2014). Research on influencing factors of tourist satisfaction of Cultural heritage Scenic spots based on Grounded Theory: A case study of tourist comments of Beijing 5A Scenic spots on Dianping. *Economic Geography*, 34(01):168–173+139.
2. Hussein, D. M. E. D. M. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University - Engineering Sciences*, 30(4), 330–338.
3. Liu D.Z (1979). Suzhou classical gardens. China Building Industry Press.
4. Liu W.L, Huang W. (2021) Sentiment clustering of landscape constituents of Liuyuan Garden based on domain dictionary. *Science Technology and Engineering*, 21(8): 3174-3179.
5. Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113.

6. Peng Y.G (1986). Analysis of Chinese classical gardens. China Building Industry Press.
7. Zhang R.Y., Wang J. N., Huo Y.D. (2021). Exploring the heritage value of Suzhou classical Gardens from the perspective of tourists – Taking Humble Administrator’s Garden as an example. China Cultural Heritage, (05):80-85.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

