Tourism Management and Technology Economy (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/tmte.2025.080130
ISSN 2616-2199 Vol. 8 Num. 1

Perceiving Tourism Image through User-Generated Short Videos—A Case Study of a Chinese National Tourism Resort

Xiaojuan Rao

School of Modern Transportation and Business, Tourism College of Zhejiang, Hangzhou, 311231, China
839287157@qq.com

Keywords: User-Generated Content (UGC), Destination Image Perception, Cognitive Destination Image, Affective Destination Image

Abstract: This study explores tourism image perception based on official Douyin User-Generated Content (UGC) short videos from the Xianghu National Tourism Resort. Using content analysis, it examines both cognitive and affective dimensions of destination image. Findings reveal that the cognitive image is dominated by traditional natural landscapes, while cultural elements are notably underrepresented. The affective image is largely associated with "pleasant" emotions. Given the interdependence between cognitive and affective images, enhancing the resort's tourism image requires shifting visitor attention from natural and architectural features toward folk customs, cultural heritage, and tourism activities. Diversifying emotional expressions is also essential to foster a more comprehensive destination image.

1. Introduction

New media technologies have diversified tourism information channels, shifting the focus from traditional to digital media. Short video platforms, characterized by strong social interaction, low entry barriers, rapid diffusion, fragmented content, and vivid visual appeal, have become increasingly popular among audiences. Chinese official statistics indicate that short videos significantly promote tourism and drive product sales. China had 940 million internet users, with an internet penetration rate of 67% by 2020. Reports further indicates that the platform hosts over 850 million monthly active users by 2020.

Users across online platforms increasingly share original content—photos, audio, videos, and text—fueling the rise of User-Generated Content (UGC). In tourism, short-video creators post original videos on UGC platforms where audiences interact through viewing, liking, commenting, saving, and sharing, signaling endorsement and engagement. UGC—especially video-based content—has become a key medium shaping destination image and influencing travel decisions. As a leading UGC short-video platform, Douyin now serves as an essential marketing tool for many destinations and scenic areas. Short videos convey destination images in a multidimensional and dynamic manner, fostering interactivity, engagement, and intimacy that allow users to experience

destinations vividly and holistically. Viewers engage with creators in comment sections, using short videos both for pre-trip planning and for sharing travel experiences. Consequently, UGC short videos have evolved into vital platforms for destination image communication, significantly shaping potential tourists' decisions.

Recently, many scenic areas have launched official accounts on short-video platforms, innovating marketing strategies with notable results. Nonetheless, substantial potential for improvement remains in short-video operations across numerous destinations. This study analyzes videos under the hashtag #123 high Xianghu# from the official Douyin (Chinese version of Tiktok) account of Xianghu National Tourism Resort (a national resort located in Hangzhou, China). Using content analysis, it explores destination image through cognitive and affective dimensions.

2. Literature Review

2.1 Destination Image

A tourism destination image is defined as the impression people hold of areas where they do not reside, highlighting it as a subjective construct shaped by external stimuli [1]. This marks one of the earliest academic contributions to destination image research. Later, it was extended by incorporating emotional dimensions, emphasizing the roles of intuition, attitudes, and cognitive evaluation. Classical theories subsequently divided destination image into two types based on perspective: the projected image (supply-side) and the perceived image (demand-side).

In the contemporary UGC-driven context, the traditional projected image has evolved into a digital image jointly constructed by destination management organizations (DMOs) and tourist-generated content. Destination image is commonly understood to consist of two interrelated dimensions—cognitive and affective—where the latter evolves from the former. The cognitive image encompasses tourists' knowledge, beliefs, and evaluations of tangible destination attributes (e.g., architecture, scenery), whereas the affective image captures their emotional responses and psychological impressions, such as excitement or nostalgia.

Owing to the multidimensional nature of tourist perception, research on tourism image has become increasingly diverse in both content and methodology. Scholars classified tourism perception studies into three levels—core factor, external factor, and connotation factor perception—identifying 14 research types and situating tourism image perception within the core factor dimensions [2].

2.2 Research on Short Videos and Tourism

Compared with text-based materials, visual media such as images and videos convey destination images more vividly and intuitively. Previous studies have mainly analyzed online reviews, travel notes, and photographs [3], while microfilms and promotional videos have also been widely explored. In recent years, short-video research has flourished, providing novel approaches and channels for constructing, enhancing, and disseminating urban and tourism images. Short videos offer distinctive perspectives for representing urban culture and landscapes, though certain limitations remain. The integration of short videos with urban image construction remains relatively new. In 2018, Xi'an collaborated with Douyin to amplify its tourism image through short-video marketing, inspiring subsequent studies on "internet-famous" cities such as Xi'an and Chongqing [4]. Scholars have since examined short-video dissemination strategies for urban image, the link between short videos and urban image construction, related challenges and countermeasures, and their positive impacts on tourism marketing, confirming Douyin's growing role in urban image communication. Overall, short videos serve as effective media for shaping and interpreting

destination images, exerting significant influence on tourists' perceptions and decision-making behavior.

Studies on destination image have primarily relied on traditional questionnaire-based methods from cognitive and perceptual perspectives. However, with the proliferation of short videos, UGC materials have become more abundant, vividly reflecting tourists' experiences and offering valuable reference points for potential visitors [5]. Such content attracts attention, fosters trust, and influences travel intentions, providing new materials and perspectives for destination image research. Consequently, analyzing tourism image through short-video content has emerged as a growing research focus.

Current studies on online destination image, primarily based on textual analysis through surveys and interviews, are relatively mature but often limited by small sample sizes and subjective biases. Research focusing specifically on UGC short videos remains limited. As vivid and direct materials for image research, short videos are generally analyzed through content or semiotic analysis, with the former providing greater precision. However, manual analysis is time-consuming and labor-intensive, and small sample sizes often restrict research scope, making video-based studies particularly challenging.

Utilizing UGC short videos can broaden the scope of destination image research, mitigate the constraints of traditional methods, and enable more comprehensive multidimensional analyses—representing a valuable methodological advancement. Many destinations and scenic areas have yet to fully optimize official short-video content strategies, resulting in a gap between market demand and practice. This study seeks to offer practical insights for improving scenic-area video operations and to explore new avenues for constructing and disseminating tourism images.

3. Research Design

3.1 Research Object and Method

The official Douyin account of *Xianghu National Tourism Resort* was launched in 2020. The hashtag campaign #123 high Xianghu# ran from May 1 to August 1, 2020, with all video creators being tourists. This study analyzes 439 UGC short videos posted under this hashtag during the campaign period as its research samples.

This study adopts content analysis, a scientific method for conducting an in-depth examination of textual content. The analysis is supported by ROST Content Mining 6, a key tool that performs word segmentation, word frequency statistics, and data visualization. The software identifies and counts all words within a document, filters the frequency results, and extracts high-frequency keywords. It then visualizes the relationships among these keywords through semantic network graphs, enabling the researcher to derive persuasive and generalizable conclusions from digital data.

3.2 Research Analysis

Drawing on cognitive-affective theory, this study focuses on video content and viewer comments to examine destination image. The portrayal of human activities, object imagery, and verbal descriptions within the videos reflects tourists' cognitive perceptions of the destination, while the comments posted by viewers capture their affective responses, representing the emotional dimension of the destination image.

3.3 Analysis of Cognitive Image in Video Content

3.3.1 High-Frequency Noun Analysis

A frame-by-frame analysis of the 439 videos identified 147 distinct nouns, appearing a total of 1,131 times. Incorporating other elements—such as hashtags, main themes, captions, narrations, and background music—resulted in a corpus of 12,175 characters representing the cognitive image. Using ROST CM6 software, word frequency analysis was conducted. Related expressions were merged: for instance, various dish names were grouped under *"cuisine"*, and different titles for games between parents and children were unified as *"parent–child games."* The analysis yielded 273 high-frequency nouns with a total occurrence of 2,061 times. The top 50 high-frequency nouns are presented in Table 1.

The results show that terms such as *"flower," "Xianghu," "grass," "bridge," "mountain," "water,"* and *"tree"* appeared most frequently. The most common person-related categories included *"child," "chef," "youth,"* and *"couple."* Tourists were frequently shown engaging in activities such as *parental–child interaction*, *flower viewing*, *dancing*, *boating*, and *weddings*. These findings suggest that the natural scenery of the Xianghu Scenic Area was a primary focus of attention. Visitors generally perceived Xianghu as a beautiful and pleasant natural environment, characterized by its public, outdoor, and free recreational spaces, making it an ideal destination for family outings and leisure activities.

Table 1 High-Frequency Nouns in Video Content (Top 50 Words)

Noun	Frequency	Noun	Frequency	Noun	Frequency	Noun	Frequency	Noun	Frequency
Lake	198	Fish	33	Hall	18	Abalone sauce	12	Summer	9
Delicacy	160	Flower sea	28	People	14	Water shield	12	Wild vegetables	9
Flowers	86	Game	26	Sunset	14	Chestnut	11	Fun	8
Bridge	75	Flower appreciation	24	B&B	14	Outdoor	10	Boating	8
Cuisine	72	Lotus flowers	23	Jadeite	14	Children's Day	10	Parental-Children games	8
Mountain	70	Fly kites	23	Taro	14	Hydrangea	9	Fish paste	8
Boat	64	Pagoda	22	Chef	13	Bamboo forest	9	Clam soup	8
Trees	46	Sky	22	Children	13	Nature	9	Mandarin fish	8
Grass	42	Sights	21	Beach	12	Dancing	9	Multisecta	8
Scenery	33	Parental-children	20	Path	12	Soup	9	Travel	7

3.3.2 Dimensional Analysis of Video Content

Given that food-related content constituted a significant portion of the depicted activities, food-related nouns were established as a separate dimension. Consequently, seven core dimensions of video content were identified: Nature, Food, Activities, People, Architecture, Time, and Culture, encompassing all 2,061 noun instances. The proportional distribution of these dimensions, derived from keyword analysis, is shown in Table 2. The Nature dimension was dominant (853 instances, 41.4%), comprising terms for natural elements like "lake," "bridge," "grass," "sea of flowers," and general scenic nouns. Within this dimension, "lake" (specifically referring to "Xianghu Lake") appeared 198 times—more than double the frequency of the second-most common term, "flower" (86 times). This indicates that "Xianghu Lake," the primary attraction, was the central focus for tourists, whose attention was significantly more oriented towards the natural landscape than towards various activities or cultural aspects.

Table 2 Statistical Distribution of Video Content Dimensions

No.	Category	Frequency	Proportion (%)	Keywords
1	Nature	853	41.4	Lake, flower, bridge, mountain, trees, grass, scenery, flower sea, lotus flowers
2	Food	577	28	Fish head, restaurant, jadeite, taro, abalone sauce, water shield, chestnut, wild vegetables, fish paste
3	Activity	295	14.3	Game, flower appreciation, fly kites, outdoor, dance, fun, travel, boating
4	People	126	6.1	Parental- Children, kids, chef, son, sisters, customer, young people, the immortal
5	Architecture	125	5.3	Restaurant, B&B, farm house, hotel, holiday resort, wooden house, Kuahuqiao Bridge, square, Jinglu
6	Time	85	4.1	Children's Day, summer, Labor's Day, May, start of summer, rain, Summer Solstice, start of autumn, start of spring
7	Culture	17	0.8	Hangzhou, Westlake, Xiangshan, Xiaoshan, capital

The *Food* dimension ranked second, accounting for 28% of total occurrences, including terms such as *fish head*, *restaurant*, *jadeite*, *taro*, and *abalone sauce*, which primarily reflect culinary ingredients and cooking techniques. This prominence is closely related to the account "Xianghu Jinglu Hotel," which contributed 118 food-related videos—representing 43.2% of all entries under the hashtag campaign. All videos produced by this account focused on food preparation and maintained a highly standardized presentation format: dish names were used as titles, high-quality food images were overlaid with large, bold yellow text indicating the dish name, and short clips—each lasting only a few dozen seconds—demonstrated the cooking process through concise narration. These videos effectively showcased the preparation methods of Xianghu's signature dishes.

The *Activities* dimension ranked third, representing 14.3% of occurrences. This dimension included terms such as *games*, *flower viewing*, *kite flying*, and *dancing*, illustrating tourists' various forms of engagement within the Xianghu scenic area. These depictions provide an authentic reflection of the leisure behaviors of visitors from different demographic groups. Frequent mentions of solar terms between *Start of Spring* and *Start of Autumn*, as well as the months of May and June, suggest that visitation to Xianghu is heavily influenced by seasonal factors—an influence closely aligned with the area's inherent natural characteristics.

The proportions of the *Nature*, *Food*, and *Activities* dimensions exhibited substantial variation, whereas the *People*, *Architecture*, and *Time* dimensions showed relatively similar proportions. The *Culture* dimension was the least represented, accounting for only 0.8% of total occurrences. It included place names such as *Hangzhou*, *West Lake*, *Xiangshan*, and *Kyoto*, which were primarily mentioned in comparative contexts with Xianghu's environment. Notably, video content referencing the world-renowned Kuahuqiao Cultural Heritage site or the Yuewang City ruins was exceedingly limited, indicating that Xianghu's cultural heritage resources were underrepresented in user-generated content.

3.3.3 Analysis of Affective Image from Video Comments

User comments, representing viewers' affective responses to the video content, were collected using Python software, yielding 3,656 comment entries totaling 35,960 characters. The ROST CM6 software was employed to analyze high-frequency words and extract adjectives conveying emotions and moods. These adjectives were regarded as affective feedback on the destination image conveyed by the videos, serving as the basis for word frequency and sentiment analysis.

Following the classification of eight affective traits proposed by Russel and Pratt, the emotional

adjectives appearing in the comments were categorized and their relative frequencies calculated. Among the top 300 high-frequency words, 51 emotional adjectives were identified, appearing a total of 735 times. As shown in Table 3, five affective traits were identified. The “Pleasant” trait (65.17%) dominated the overall sentiment, indicating that most commentators expressed positive emotions toward the videos. The adjectives “beautiful,” “delicious,” and “good-looking” were the most frequent, appearing 147, 103, and 63 times, respectively. “Beautiful” was mainly associated with natural scenery, “delicious” referred to food, and “good-looking” was used to describe both culinary and scenic elements.

In addition, “Arousing” (10.07%) and “Exciting” (14.15%) affective tendencies were also evident among the comments. During text preprocessing, common numerical slang expressions on short video platforms—such as “66” and “666” (means awesome) appearing 45 times in total. These expressions were mainly associated with videos showcasing the exquisite culinary skills of master chefs and conveyed meanings such as “excellent” or “amazing.” Functioning as textual equivalents of strong approval, they carry a deeper sense of endorsement than simply clicking a “like” icon, reflecting the straightforward and emphatic attitude of online commentators. The “Distressing” trait accounted for only 4.21% of emotional expressions, while the remaining affective tendencies can be interpreted as collectively forming a positive affective image of Xianghu.

Table 3 Statistical Table of Affective Traits

No.	Trait	Emotional Adjectives (with frequency)	Proportion (%)
1	Arousing	666 (45), impressive (25), extraordinary (4)	10.07
2	Exciting	Perfect (20), Hardworking (17), excellent (17), the best (16), great (8), exquisite (8), successful (7), solid (7), classic (4)	14.15
3	Pleasant	Beautiful (147), delicious (103), good-looking (63), yummy (19), thoughtful (16), happy (13), conscientious (12), detailed (8), pleased (7), stunning (6)	65.17
4	Relaxing	Free (11), simple (8), ordinary (7), refreshing (7), carefree (4), comfortable (4)	6.4
5	Distressing	Not enough (16), tiring (16), pitiful (5), outdated (4)	4.21

4. Conclusions

Based on an analysis of User-Generated Content (UGC) short videos, this study explored tourism image perception of Xianghu Tourism Resort. Findings indicate the cognitive image remains predominantly shaped by natural landscapes, with elements like "lake," "bridge," "flower sea," and "trees" frequently featured. These spatial attributes constitute the largest portion of Xianghu's image, while scenic and leisure activity content also features prominently. Collectively, these elements form Xianghu's core cognitive image, though cultural aspects remain underdeveloped.

The affective image is dominated by "Pleasant" emotions (65.17%), supplemented by "Arousing" (10.07%) and "Exciting" (14.15%) tendencies, with positive elements exceeding 95%. Affective responses are closely linked to cognitive perceptions: nature and gastronomy evoke reactions like "beautiful" and "delicious," while culinary skill videos generate "awesome" and "amazing" responses. Overall, Xianghu's affective image demonstrates positive, rich, and layered characteristics.

Recommendations for improving video marketing include:

Firstly, implement All-for-One Tourism through cross-industry integration and regional management to foster organic content dissemination. Tourism promotion should diversify themes, as demonstrated by Songcheng Performing Arts' time-travel campaigns attracting over ten million

participants. Marketing should leverage trending hashtags like "Intangible Cultural Heritage Partner" (2.28B views) and "Vlog Food" (26.67B views) to shift focus toward cultural content and encourage UGC creation.

Secondly, utilize official accounts like "Hangzhou Release" to construct Xianghu's multidimensional image, integrating cultural narratives with urban life to showcase Hangzhou's fusion of tradition and modernity. Food content should associate with broad hashtags like #IAmAFoodCreator (16.58B views) rather than limited-traffic #XianghuFood (43.36M views) to enhance visibility.

Lastly, leverage foreign residents' perspectives to strengthen Xianghu's global image. The Russian-run account "Demiri" gained million-plus followers by sharing Chinese culture through #ZhejiangGoodRecommender. Such cultural intermediaries can highlight core symbols like Kuahuqiao Heritage and "Xianghu Dialogue" as international hallmarks, particularly in the pre-Asian Games context.

Acknowledgements

This paper was supported by Zhejiang Provincial Education and Teaching Reform Project (NO.2021YB25), Regular Planning Project of Philosophy and Social Sciences in Hangzhou (NO.G21YD005), and 2020 School-Level Regular Project of Tourism College of Zhejiang (NO.2020ZD01).

References

[1] Hunt J D. (1975). Image as a factor in tourism development. Journal of Travel Research, 13(3), 1-7.
[2] Han X, Liu A L. (2019). Research content and evaluation methods of tourism perception. Tourism Tribune, 34(04), 106-118.
[3] Deng N, Zhong L N, Li H. (2018). Destination image perception based on UGC image metadata—A case study of Beijing. Tourism Tribune, 33(01), 53-62.
[4] Li J Y, Yan Y, Hu L L. (2019). Short video social perspective: Destination image and tourists' behavioral intention—A case study of Xi'an. Journal of Ningxia University (Natural Science Edition), 40(02), 176-184+190.
[5] Rao X J. (2019). Content analysis of tourism image of Songcheng Scenic Area, a benchmark of night cultural and tourism economy. Journal of Hanjiang Normal University, 39(06), 57-60.