Environmental Challenges 5 (2021) 100360

Contents lists available at ScienceDirect

Environmental Challenges

journal homepage: www.elsevier.com/locate/envc

Understanding the visual image of Kailash Sacred Landscape through geo-tagged landscape photos mapping

Yucheng Zhang[a], Jie He[b,*]

[a] School of Architecture, Tianjin University, No.92 Weijin Road, Nankai District, Tianjin, 300072, PR China
[b] School of Architecture, Harbin Institute of Technology (Shenzhen), University Town of Shenzhen, Nanshan District, Shenzhen, Guangdong, 518055, PR China

A R T I C L E I N F O

Keywords:
Computer vision
Kailash Sacred Landscape
Landscape image
UCG geo-tagged photos
Viewshed analysis
Visual content analysis

A B S T R A C T

Kailash is one of the most important sacred landscape in Trans-Himalayan region and worshiped by multiple nations and religions. The visual construction of this sacred landscape is a complex process of cultural practice, and even in the era of Internet and big data, the burgeoning photo sharing websites, social media have further influenced people's gaze and imagination of Kailash. The research paper focuses on people's visual perception of Kailash Sacred Landscape and proposed a hypothesis that the images of Kailash become a product of viewer's selection of iconic symbols from landscape and then be reconstructed through visual media. The research mainly includes two parts. The first part focuses on the GIS spatial analysis of the complex visual landscape of Kailash. Viewshed analysis reveals the visual prominence of Kailash in the whole region, and also the change in people's visual perception of this mountain during the kora trek. The second part focuses on of the geo-tagged photos collected form social media. Visual content analysis based on computer vision is adopted to recognize items from collected photos. Most visual elements can be integrated into a higher-level set of 8 visual elements. According to the ratio of each element, hierarchical clustering is adopted to describe the similarity between each photo. The dendrogram of clustering result shows all the photos can be generally divided into 9 groups, which reflects different paradigms of people's visual expression of Kailash. In conclusion, this research reveals how Kailash is represented through filtered images in a web-based context. Also, it explores and extends the applications of GIS and AI technology in cultural landscape research.

1. Introduction

Landscape as an important concept of cultural geography draws people's attention to the visual aspects of the relations between the modes of human occupation and construction of spaces (Cosgrove, 2003). Study on landscape has had a long history from the perspective of visual media. Back in 17th centuries, the Dutch landscape paintings were seen as "a picture representing a view of natural inland scenery", which reflect a nascent form of national identity (Olwig, 1996). But not confined to landscape paintings, a series of visual media like drawings, photos, advertisement and movies are also the media of visual representation of landscape. Especially after the "visual turn" of geography (Thornes, 2004), visual media representing and describing the landscape are no longer regarded as neutral, but containing a series of dynamic cultural and social relations (Wang et al., 2017). "Landscape" can be considered as a skillful and cultivated way of visually forming and representing a physical environment, fusing "reality" with images (Urry and Larsen, 2011). The representation of places in the image is conducted through "spatial selection", a process during which people select a specific part of landscape to represent the information and characteristics of a whole region, using "difference" to create their own geography (Van Gorp and Béneker, 2007). Those particular views considered photogenic or iconic, such as landscapes, people, buildings, are sought out by people and visually reproduced in the images (Garrod, 2009; Balomenou and Garrod, 2019).

As one kind of the largest media of human civilization, photos is an important material for us to study the characteristics and dynamic change of landscape (Liu and Wang, 2019). Especially under the influence of Internet and information industry development, spatial relations are undergoing a fierce evolution (Starrs, 1997). The popularization of Internet media has greatly expanded people's ways of landscape perception and made it break through the boundary of traditional physical space. Meanwhile, with the rapid development of various social media and image sharing platforms, people have entered an era of image. The available image data is increasing at an unprecedented speed. Photo-sharing websites like Flickr, Pinterest and Instagram have stored hundreds of millions of pictures of their users. In this context, many users begin to publish their original content through the network platform and the concept of "UGC (user-generated content)" is blooming (Chen and

* Corresponding author at: Harbin Institute of Technology (Shenzhen), China.
E-mail address: hejie2021@hit.edu.cn (J. He).

https://doi.org/10.1016/j.envc.2021.100360
Received 17 July 2021; Received in revised form 16 October 2021; Accepted 1 November 2021

Zhang, 2015). Photos are among one of the main information carriers for UGC due to its straightforward way of information transmission. It can be seen as a record of images that people, especially tourists have gazed on (Urry and Larsen, 2011). Photos on social media accompanied with texts share a way for people to construct or reconstruct their experiences of the places (Lo and McKercher, 2015). Especially UGC geo-tagged photos contain both spatial and visual information that can be considered to an indicator of cultural ecosystem service usage (Richards and Friess, 2015; Richards and Tunçer, 2018). Therefore, UGC photos are an important medium for transmitting people' perception of landscape, which has high reference value for the reconstruction and understanding of landscape images.

Mt. Kailash(冈仁波齐峰 in Chinese), at 6638m above sea level, is a peak located in the north of Darchen of Burang County, Ngari Prefecture of Tibet Autonomous Region of China, and also the main peak of the Kailash Range. Looking like a "pyramid" with four symmetric sides, the shape of Kailash differs a lot from the nearby mountain peaks. Being considered as the most sacred mountain in Tibet, Kailash is a mountain full of culture and history. Its influence has also expended to a vast area across China, India and Nepal, even spreading all over the world. The formation of Kailash Sacred Landscape is a complex process of cultural practice. Fundamentally, Kailash is considered as a sacred site in four religions: Tibetan Buddhism, Bön, Hinduism and Jainism. Every year tens of thousands of believers make a pilgrimage to Kailash in the form of Kora, a holy ritual that circumambulate the sacred mountain on foot. Meanwhile, the modern "rediscovery" of Tibet in the 20th century transformed this local sacred site into a world-famous mountain (McKay, 2015). Nowadays in the era of Internet and big data, the burgeoning photo sharing websites, social media have further influenced people's gaze and imagination of this sacred mountain. A series of media like travel guides, movies, documentaries have extremely expended people's ways of landscape perception, which transcends the traditional space-time distance and becomes an important way to convey and express the image of Kailash.

As one of most sacred mountains in the world, Kailash is famous for its lofty and impressive peak which has a strong power to raise people's sense of sacred. Historically, the image of this mountain has been depicted by all sorts of visual media from believers' religious artwork to travelers' photos. Especially, social media disseminated through the Internet further enables people to appreciate the mountain landscapes without the effort of a real trek, which provides new method for inferring the visual image of this ancient sacred landscape. This paper focuses on the visual construction of Kailash Sacred Landscape in a web-based context, supposing that the generation of Kailash image is through people's selection of symbolic elements from the landscape and then reconstructed through visual media. By using virtual landscape model and geo-tagged photos as the main source of data, our research aim includes two parts: (1) to explore the relationship between the visualscape of Kailash and spatial distribution of photos, and (2) to find the most representative elements of the landscape and the paradigms within the photos. Meanwhile, the effectiveness of applying GIS spatial analysis and computer vision technology in landscape research is also discussed.

2. Material and method

To conduct this research, the virtual landscape model of Kailash Sacred Landscape is reconstructed based on the GIS platform. Geo-tagged photos were collected for mapping the spatial distribution of visual elements. Besides the viewshed analysis, a combination of visual content analysis and computer vision is also used to identify the visual elements from thousands of photos.

2.1. Research area

Just like its universal cultural impact, the area related to Kailash is not confined to a single peak. According to the world heritage tentative list of China (UNESCO World Heritage Centre, 2017), "Scenic and historic area of Sacred Mountains and Lakes" is a region includes two sacred mountains: Gang Rinpoche (Kailash) and Naimona'nyi, two holy lakes: Manasarovar and Lhanag-tso, and eight temples around Manasarovar. While in the transboundary project of The International Centre for Integrated Mountain Development (ICIMOD), "Kailash Sacred Landscape" is defined as a vast landscape area of over 31,000 km^2 across three countries including China, India and Nepal (Zomer and Oli, 2011).

Regardless of the wide influenced region of Kailash, this paper mainly focuses on the people's visual perception of the peak of Kailash. Therefore, our overall research area is the same as the Chinese part of "Kailash Sacred Landscape Conservation and Development Initiative (KSLCDI)" hold by ICIMOD. Especially, a 10*10km square around the peak of Kailash is considered as the core research area (Fig. 1), which totally covers the outer path of kora and its surrounding area.

2.2. Data

The research data mainly comes from two sources. One part of data combines digital elevation model (DEM), hydrological data and road data to reconstruct a virtual landscape model of Kailash Sacred Landscape. The accuracy of DEM data is 90m, which is precise enough for describing the huge mountains in the vast region. The 1:250,000 scale hydrological and road data is provided by National Catalogue Service for Geographic Information (2021) of China.

Another part of data is geo-tagged photos collected from two famous online tourist platforms (www.foooooot.com and www.2bulu.com) in China. We use the Chinese keywords including "冈仁波齐" (Kailash) and "普兰县" (Burang County) to locate and search for the related photos on the two websites. Photos are downloaded through Web crawler technology, which can automatically collect information from the Internet by simulating people's browsing behavior. With GPS chips inside cameras or smartphones, geographical information has been documented into these photos' EXIF (Exchangeable Image File Format). Therefore, latitude and longitude information of these photos can be automatically extracted (Jiang et al, 2013). Since GPS information inevitably has some errors due to some confounding problems, we further manually filter those photos following three rules. (1) We removed all photos that were taken outside our research area. (2) We identified and removed duplicate photos all attributes of which are the same. (3) We filter those photos the theme of which is unrelated to landscape. Finally, 8,092 photos in total are collected, all of which distribute in the core research area. These photos with geo-tag enable us to link the content of photo with real visual environment, facilitating the obtaining of spatial distribution characteristics.

These photos are mainly taken and uploaded by tourists and believers from China, which can be seen as interpretation of the image of Kailash Sacred Landscape from perspectives of the contemporary society of China. Though other photo sharing websites like Flickr also provides numerous photos about Kailash in global range, but very few of these photos contain geographical information which enable us to trace the spots where the images were shot.

2.3. Method

2.3.1. Visualscape analysis

The visualscape can be considered as "the spatial representation of any visual property generated by, or associated with, a spatial configuration" (Llobera, 2003). This concept is quite useful for us to understand the special visual property of Kailash. All the visualscape analysis in this research is based on viewshed analysis, which is used to divide the area into visible and not visible by calculating which locations in the DEM can be connected through a straight line to the given viewpoint within any specific distance (Llobera, 2003). The extended viewshed analysis tools including total viewshed and visual exposure analysis are also used.

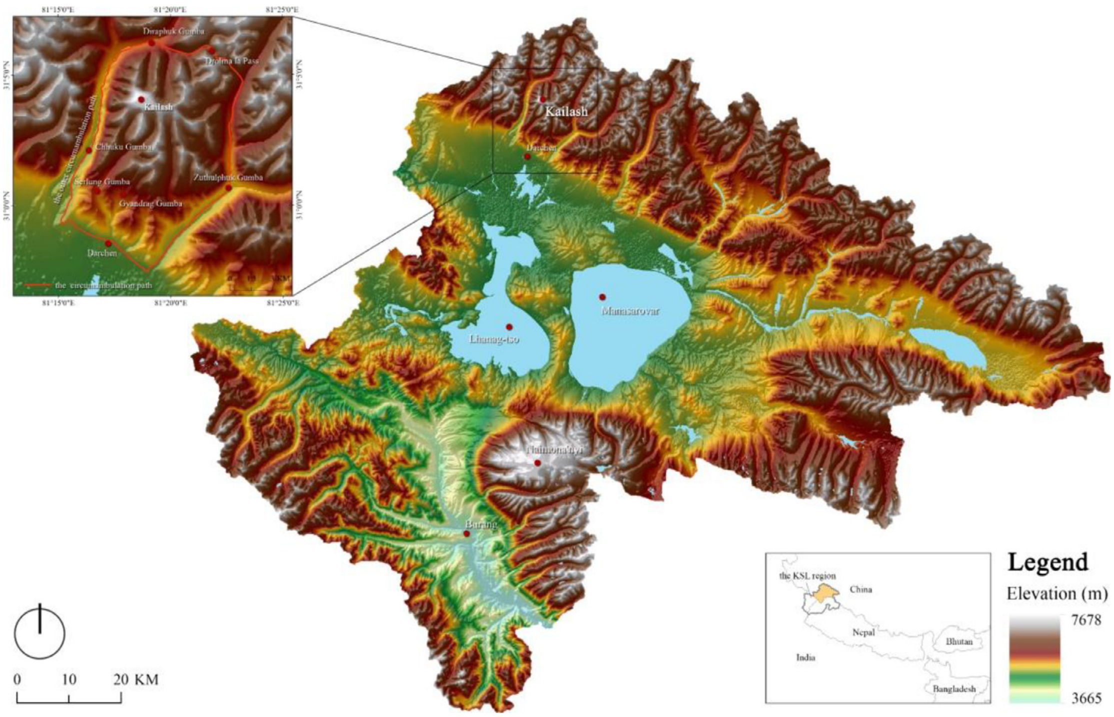

Fig. 1. The research area.

Total viewshed analysis which can provide an entire terrain with a first description of its visual structure (Llobera, 2003) is used to reveal those things that looks prominent in the landscape. Considering the harsh environment of Kailash where most areas are inaccessible to human, we firstly defined those areas where people could possibly reach on foot. The walking zone was calculated according to Tobler's hiking function (Tobler, 1993), considering the slope of terrain and inaccessible waters. We only located the observation points in each cell of DEM which could be reach within 1hr by walking from the road. Then, we calculated the viewshed of each view point and finally added them up to achieve the total viewshed value. Those places with high value tend to have higher possibility to be noticed by people in the landscape.

Another kind of visualscape called visual exposure is used to describe the visible portion of whatever concerned of the investigation, and it can be generated by visual angles (Llobera, 2003). The value of visual exposure was described by calculating the visual angel from the given view point to the target. In this research, to describe the visual attraction of Kailash to other areas, we firstly conducted viewshed analysis to define areas where can see the peak of Kailash. Then, the visual angels between each view point within the visible areas and the peak were calculated in order to represent the visual exposure of Kailash. The vector from low value area to high value area can be seen as the possible direction of people's line of sight when attracted by Kailash.

2.3.2. Visual content analysis based on computer vision

Visual content analysis is a method originated from the interpretation of written and spoken texts, counting and then analyzing the frequencies of certain visual elements in a well-defined collection of images (Rose, 2016). This attribute-based method helps identify the main items in the photos and record their frequencies, co-occurrence, clustering, and other related issues in a quantitative way (Stepchenkova and Zhan, 2013). Though with a relatively well-established theoretical framework, the traditional human-based content analysis of images is quite time consuming, and also suffers from the instability and irregularity of artificial coding (Zhang et al., 2019). To deal with thousands of digital images, a variation called cultural analytics was invented, which is an "automatic computer-based method to describe large numbers of cultural objects quantitatively" (Manovich and Douglass, 2011). Especially, with the development of computer vision and machine learning technology, the intelligent processing of big data and full use of content and value of picture has become possible. Semantic segmentation, a kind of computer vision technology based on deep learning is quite useful in extracting elements from images. It can identify every pixel of the image into different elements, outputting their location and percentage. Meanwhile, the large-scale image datasets like Cityscapes (Cordts et al., 2016) and ADE20k (Zhou et al., 2017, 2019) provide sufficient data for researchers to train the model and identify the elements of the image accurately.

In this research, visual content analysis based on semantic segmentation is adopted to extract visual elements from geo-tagged photos. The DeepLabv3+ model with Xception as network backbone pretrained on ADE20k dataset provided by PixelLib in Python (PixelLib, 2021) is used to realize image semantic segmentation, which can identify at most 150 classes of objects from images. Comparing with other network, DeepLabv3+ has been proved to be a faster and stronger encoder-decoder network and thus has more effective performance (Chen et al., 2018). The pixel-wise accuracy of this model is 82.52% on the validation set. Therefore, it can relatively accurately segment different visual elements from the photos, as well as their composition and proportion. Four examples of the outputs are shown in Fig. 2.

2.3.3. Methods for spatial and statistical analysis

In this research, fishnet analysis was used for geo-visualization and spatial analysis, and hierarchical clustering was used for statistical analysis.

A fishnet in ArcGIS was created to divide the core research area into 100m square grids. The number of photos in each grid was counted after all geo-tagged photos were located according to their longitude and latitude attributes. The average value of different photos' attributes within the grid was calculated to represent the corresponding grid, through which the spatial distribution and characteristics of photos could be displayed. Further, the spatial distribution characteristics of image content were discussed from two aspects: diversity and representativeness.

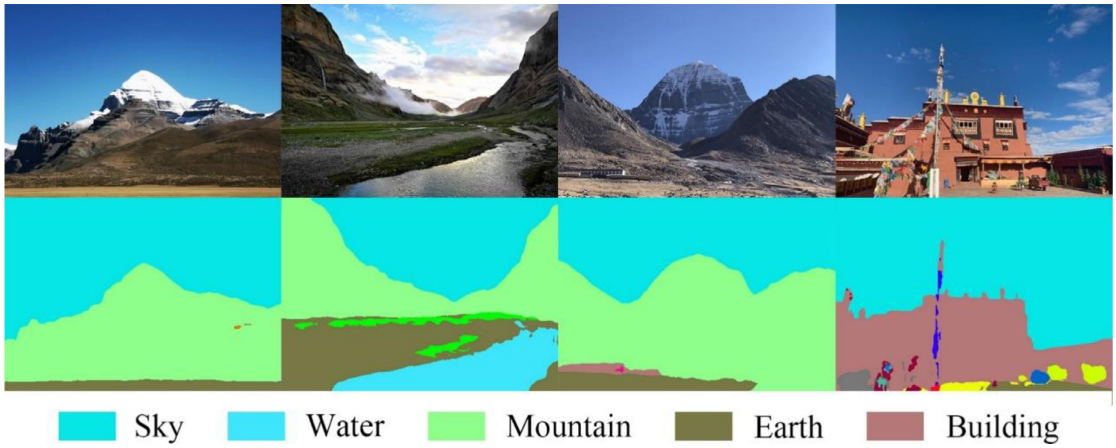

Fig. 2. Examples of the outputs of the image semantic segmentation.

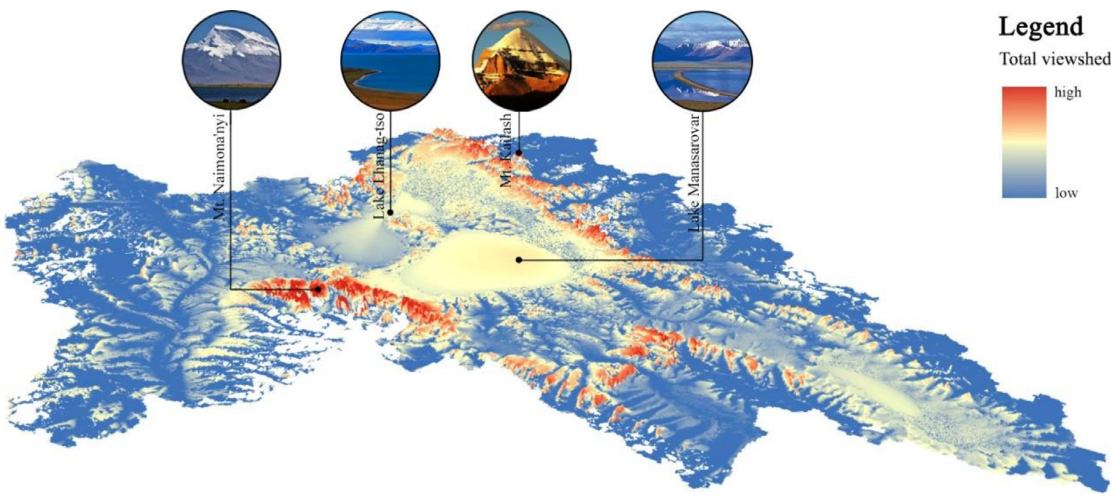

Fig. 3. Total viewshed analysis of Kailash (3D).

Diversity can reflect the number of elements people want to express in a single photo. A photo with high diversity tends to be more comprehensive and complex, while a photo with less diversity tends to be focused and simple. To describe the diversity of elements in the photos, we adopt Simpson's diversity index, which is initially used to describe the biodiversity and later also used in describing landscape diversity. The formula for calculation is as follow (Eq. 1) and the average diversity within a grid is calculated:

$$D = 1 - \sum_{i=1}^{S} P_i^2 \tag{1}$$

Where D is the diversity of elements in the photo, S is the number of all kinds of visual elements, Pi is the proportion of element i's pixels in the photo.

Representativeness is defined to reveal the most common element in a given area. To find the representative of a certain area, each photo is represented by a visual element that accounts for the largest proportion in the image, then the element that represents the most photos in the given area is regarded as the representative of the area. This representative can reflect the element within the landscape that leave the deepest impression to people.

Apart from the computing of some indices, hierarchical clustering also is applied in the research. Hierarchical clustering is an unsupervised machine learning method, which successively partitions the dataset into different hierarchies according to the similarity or distance between network nodes, and then form a dendrogram (a tree-like graph hierarchical structure). In this research, a hierarchical clustering algorithm using Ward's distance is implemented in the SciPy in Python (The SciPy community, 2021), considering the photos' latitude, longitude and proportion of different elements in the images. With this method, the similarity and difference of photos in each grid could be explored.

3. Result

3.1. The visualscape of Kailash

The visual environment of Kailash is the basis of people's perception of the landscape. To understand how people seek out views from the environment, the GIS spatial analysis is used to reveal the relationship between the geo-location of photos and the visualscape of Kailash.

Fig. 3 shows the total viewshed of the whole research area in 3D form. The deeper the red, the more visibility the area has. It can be seen that the peak of Kailash and Naimona'nyi are of the relatively high visual prominence among the landscape, which endow themselves with people's high attention and the sense of sacred. With two lakes sandwiched in between, the basic landscape pattern also forms which can be described "two peaks perching to the north and south, two lakes reflecting with each other in the east and west" (UNESCO World Heritage Centre, 2017).

When it comes to the core research area, Fig. 4 shows that Kailash as a lone mountain peak, compared with other mountain ranges also with high visual prominence, stands out independently from its surround-

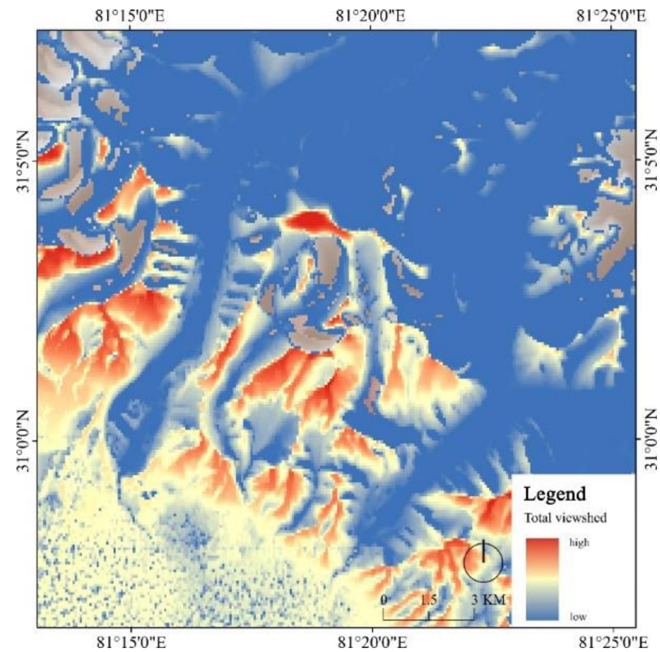

Fig. 4. Total viewshed analysis of the core research area.

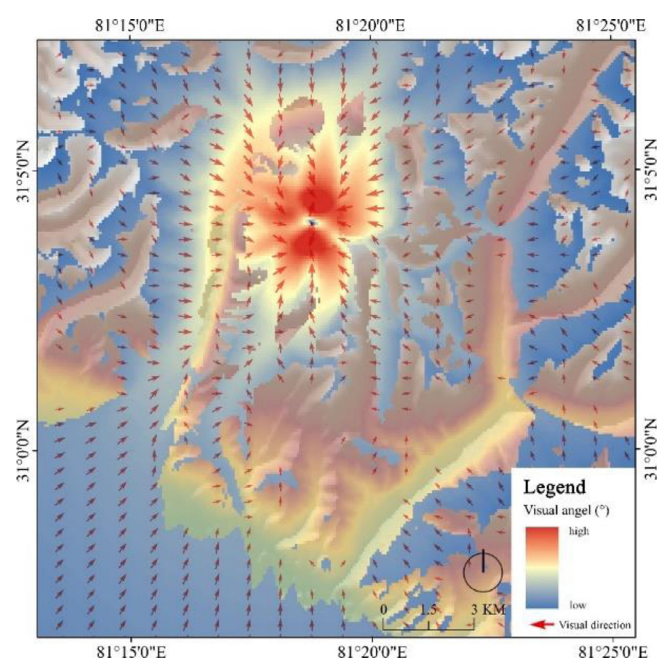

Fig. 5. Visual exposure and direction of maximum gain.

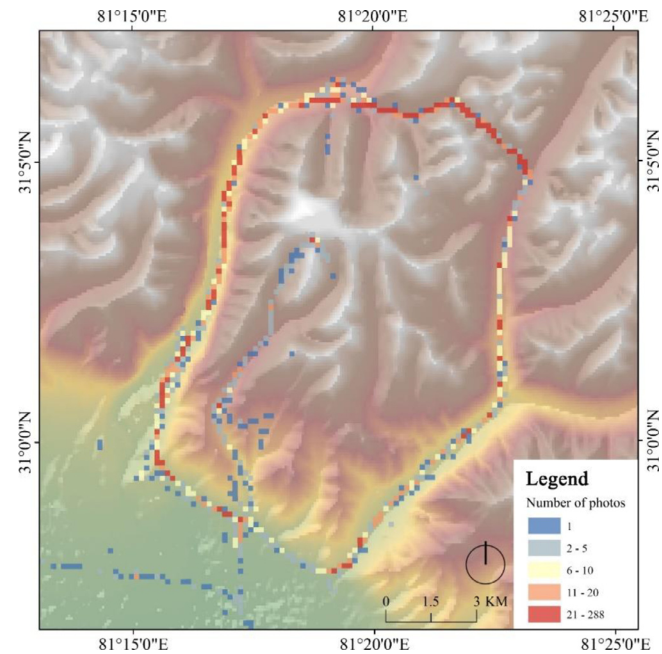

Fig. 6. The spatial distribution of photos.

ing peaks. This may explain why Kailash becomes the most prominent mountain in the range, though it is not the highest mountain in the area. Taking the peak of Kailash as the observed target, its visual exposure is revealed in Fig. 5. The arrows in the figure shows the direction in which the visual exposure value increases, reflecting how people's visual attention gradually being attracting to the peak. Stretching out from the peak, the surrounding valleys form several natural visual corridors, which draw people's attention from many directions, especially to the south and north face of the mountain. These visual corridors have become conventional scenic spots for people to appreciate the mountain and lead to the generation of some fixed images of Kailash.

3.2. Spatial distribution of geo-tagged photos

Based on the understanding of Kailash's visualscape, we use the spatial distribution of photos to verify how people interact with the landscape and where they consider as iconic and representative. Fig. 6 shows photos are generally distributed along the route of kora, the outer path in particular. Considering the clockwise circumambulation, the number of photos gradually increases when people start their trek from Darchen. The image grids with largest number of photos appeared around the northern part of the outer path, which reflects these landscapes are mostly selected to represent Kailash Sacred Landscape.

The kernel density of photos was calculated in ArcGIS. The relationship between the change of photo density and the visualscape of Kailash along the outer path is revealed in Fig. 7. Mostly, the photo density increases in those areas when the peak is visible and of high visual exposure, which means people are more willing to take photos when they are attracted by the peak. But when the peak is less visible, people are less likely to take the scenes as the image of Kailash. Also, the photo density increases when people see some important landmarks during the trek. In particular, the place where people take the most photos have nothing to do with the visual exposure of Kailash. The image grid with the greatest number of photos occurs when crossing Drolma la pass, which is the most difficult and steepest part of the trek. This phenomenon shows the image of Kailash is not only confined to the sacred peak, but also the arduous trek.

The content of photos reflects what elements are spatially selected from the landscape to represent the whole region. The computer vision and statistics method are adopted to facilitate the information extraction from photos. Focusing on the content not only enables us to know spatial arrangement of the selected elements, but also the meaning and mechanism behind.

3.3. Visual elements of photos

Using the image semantic segmentation, we classified each pixel of photos into different objects in the picture and then computed the frequency for each of the recognized items. Fig. 8 shows the top 30 items extracted from the photos, with other items together categorized as one group. The representation of items shows a strong accumulation charac-

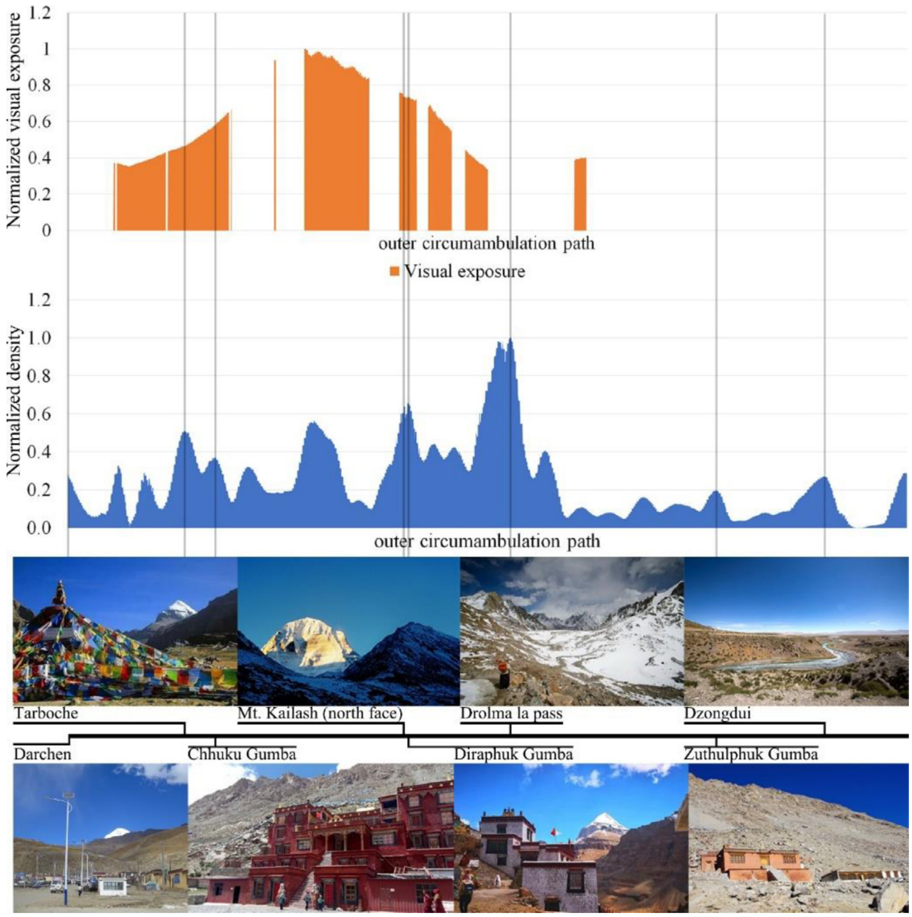

Fig. 7. The change of photo density and visual exposure during the kora trek.

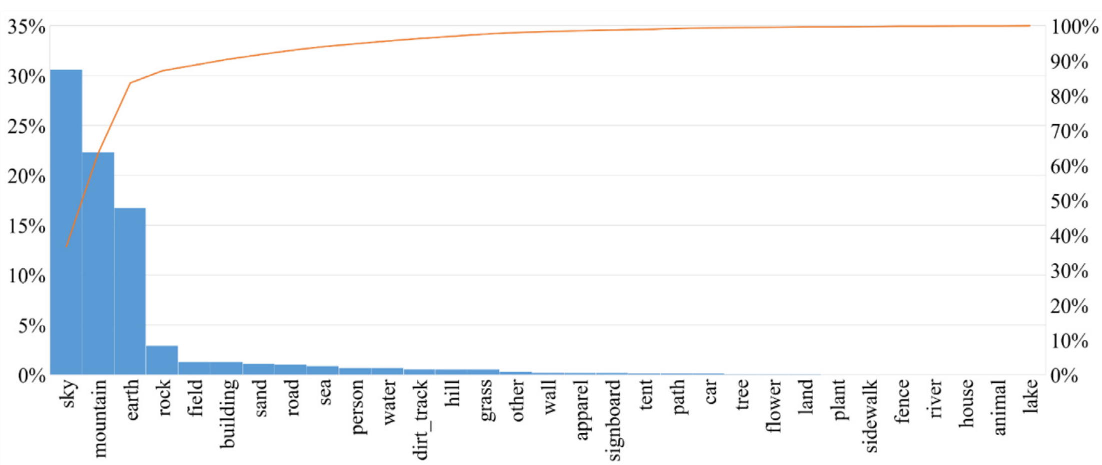

Fig. 8. The average proportion of recognized items appeared in the collected photos.

teristic. Natural items like "sky" "mountain" "earth" take the dominant place in most photos, composing the initial impression of the landscape. While some human-related items like "person" "building" "signboard" also show in some photos, which reflect the participation of human in the construction of culture landscape.

Since 150 items are too fragmented to have a general analysis of the image content, we selected and amalgamated the top 30 items into 8 visual elements according to the similarity of visual perception (Table 1), which could reflect most types of visual perception to the landscape. Fig. 9 shows the statistical differences in perception types. Still, "mountain" and "sky" generally are the top 2 elements appeared in most photos, with average proportion of 30.23% and 28.00% respectively, and "land" is the third (18.56%). These three elements form the main part or background of most photos. While compared to these three elements, the other elements occupy much less space in the photos, none of which the average proportion is beyond 2%.

To understand how people select these visual elements from the landscape and form their own experiences of the places, we need to know their spatial distribution characteristics.

The spatial distribution of the average diversity and representative element of photos are shown in Figs. 10 and 11. When combine the result with the analysis of visual exposure (Fig. 5), the relationship be-

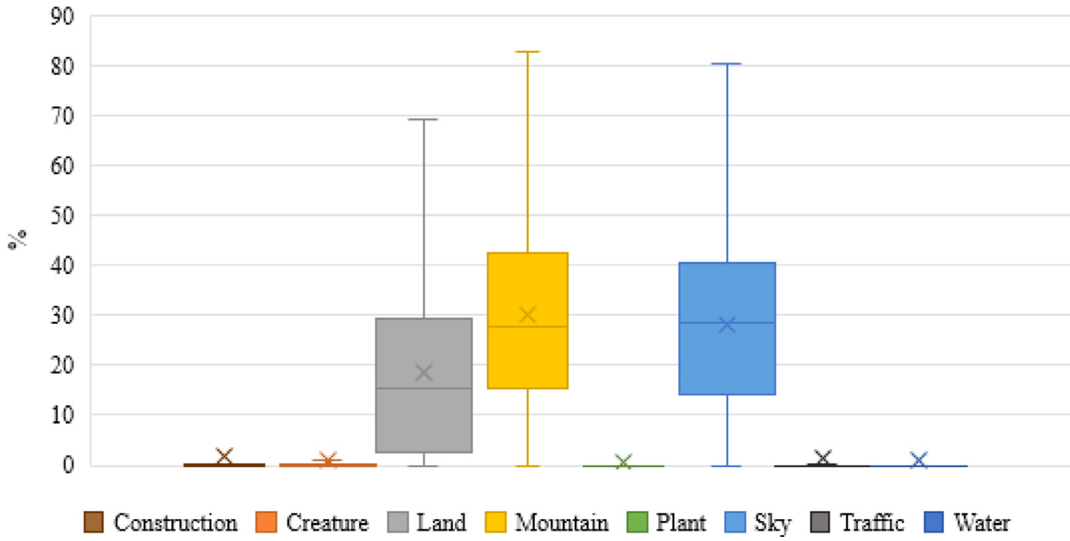

Fig. 9. The average proportion of visual elements.

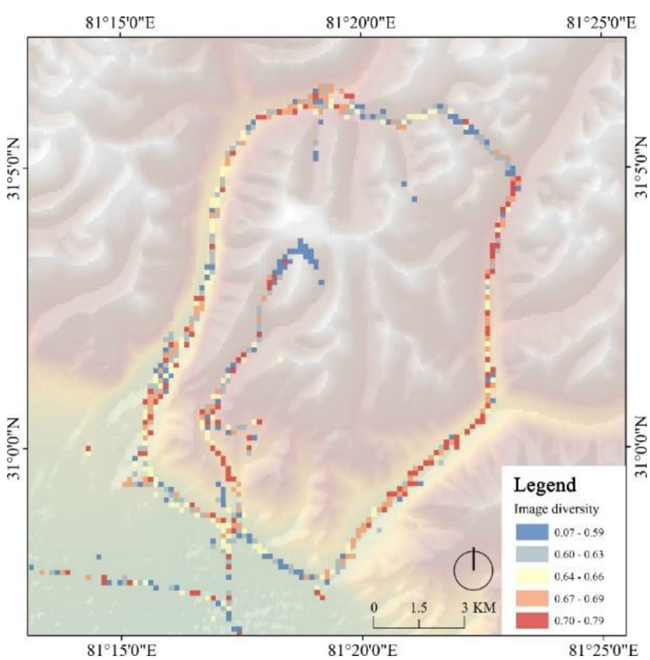

Fig. 10. Image diversity

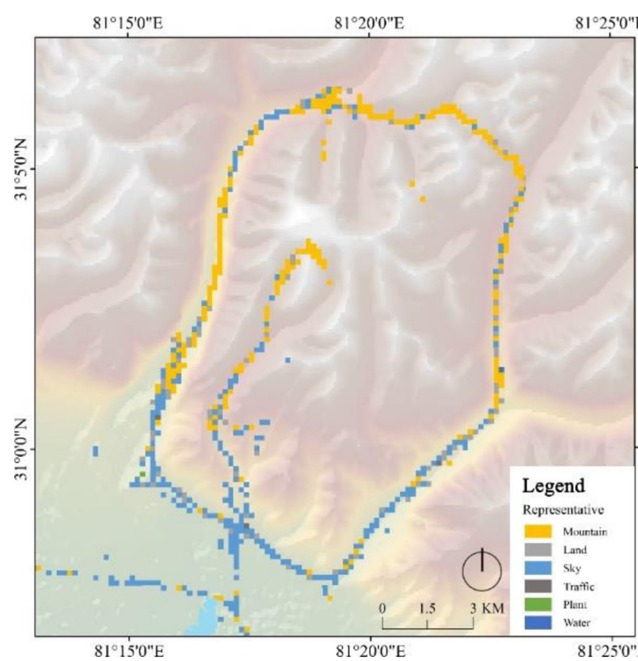

Fig. 11. Representative elements

Table 1
8 kinds of visual elements developed from top 30 recognized items.

Visual element	Items
Construction	building; tent; signboard; wall; house; fence
Creature	person; apparel; animal
Land	earth; field; sand; land
Mountain	mountain; rock; hill
Plant	grass; flower; tree; plant
Sky	sky
Traffic	car; dirt track; road; sidewalk; path
Water	water; sea; river

tween the image content and the visibility of Kailash can be revealed. In those area where people's attention is strongly attracted by the peak of Kailash, the content of image tends to have less diversity and "mountain" is the most common element. This shows a convergent inclination in people's expression of the image of Kailash when they can see the peak. These photos mainly focus on describing a pure and huge mountain peak, which accounts for the most part of the picture. While in those areas that the peak of Kailash is vague or even invisible, the diversity of image content is higher and instead "sky" becomes the most element, which indicates people's different and diversified in understanding of other areas.

3.4. Paradigms of photos

Elements in the photos are spatially selected to create a regional geography (Van Gorp and Béneker, 2007). When photos with similar elements accumulate, a "style" or "paradigm" will form to stress the unique points of the landscape (Dai and Chen, 2010). The concept of "paradigm" can benefit our understanding of the generation of Kailash image, through which some images gain the dominate position and exclude other forms of representation. We consider a group of photos with

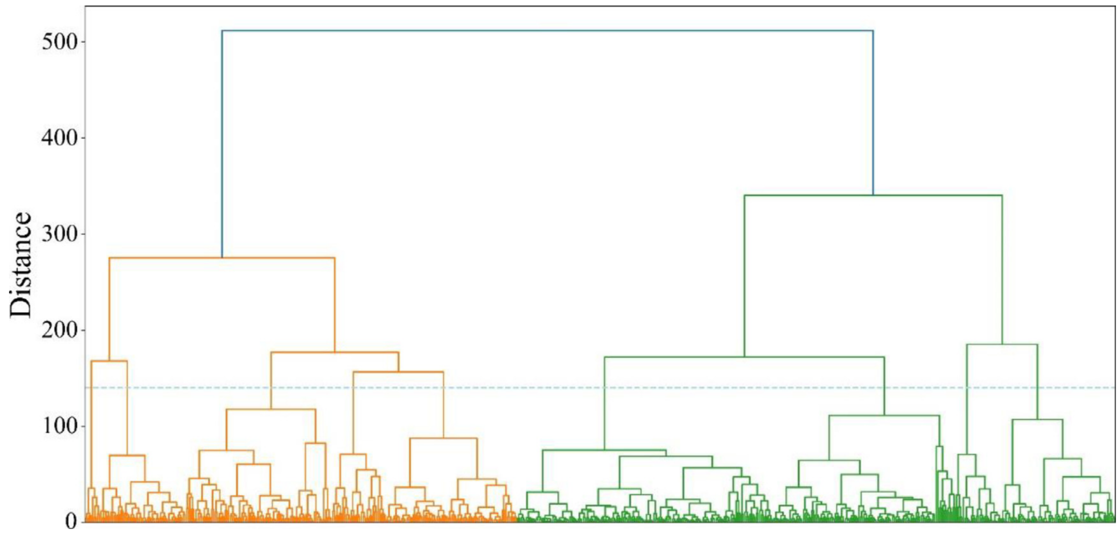

Fig. 12. The dendrogram of hierarchical clustering

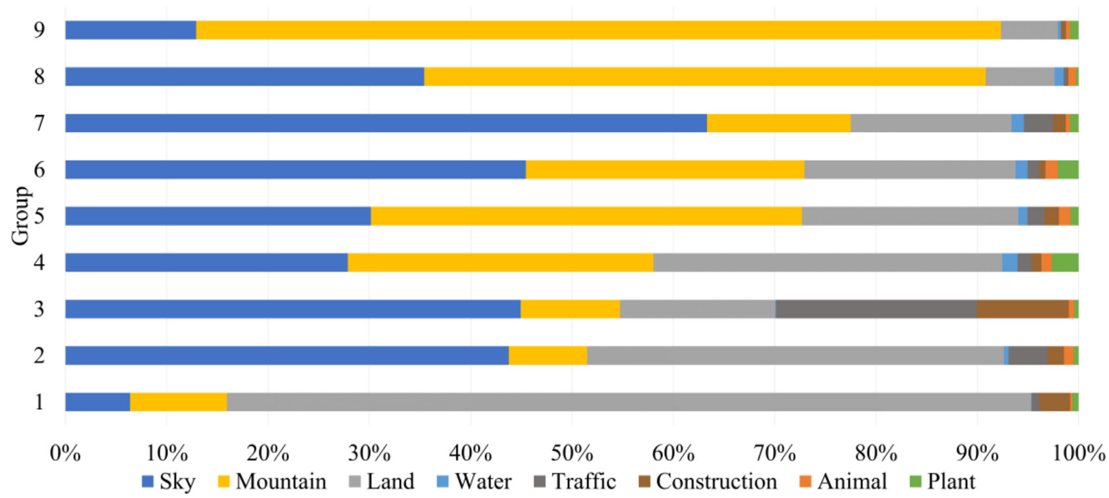

Fig. 13. The average proportion of each element appeared in different paradigms.

high similarity in their visual elements as a "paradigm". A hierarchical clustering algorithm is applied to group the photos according to the average proportion of different visual elements, as well as photos' latitude and longitude information of each grid. Fig. 12 shows the dendrogram of clustering results of all image grids, which hierarchically describes their similarity interrelation. The height of the tree (vertical axis) refers to the relative distance between different clusters. The further the distance is, the larger the characteristics difference between clusters is. To achieve clusters that are both easy to distinguish and interpret, we choose to cut the tree at the height at 140. The photos are generally divided into 9 major defined groups. The clustering center of each group can be considered as symbolic paradigms of Kailash Sacred Landscape.

To attach a meaning to each of the clusters, we calculated each group's average proportion of visual elements (Fig. 13), from which an initial impression of each paradigm could be achieved. The sequence of these groups is ranked according to the relative proportions of "land" "mountain" and "sky", which are the three dominant elements in most groups. Some extreme situations focus on only one kind of visual elements, like group 1, group 7 and group 9, elements like "land" "mountain" and "sky" take the most part of the photos. While the proportion of these three elements in other groups follow a rule of "as one falls, another rises". Such paradigms mainly focus on depicting three scenes during the kora, includes the appearance of mountain, the spacious valley, and the sinuous mountain road. Since "land" "mountain" and "sky" are the three elements with different height distribution, their relative proportions partly reflect visual angles and the spaciousness of vision, which lead to a shift in paradigms. Compared to other groups, group 3 specially focus on the human-related objects like "construction" and "traffic".

Fig. 14 shows how these paradigms are distributed, with some typical photos presented as well. The result shows when people are outside the Kailash valley, they tend to follow those paradigms with low visual angle. When people enter the valley and start their trek, however, the visual angle gradually increases and reaches the top when crossing Drolma la pass, then gradually falls down during the second half. Since people often use upward visual angles to show their worship to the observed, the fluctuation of visual angles may reflect the uplifting experience of people's spirit during the trek. When reach the north face of Kailash, people's admiration to the mountain also reaches its climax.

4. Discussion and conclusion

The visual construction of Kailash is a complex process with several clues from the numerous perspectives including religion, tourists, au-

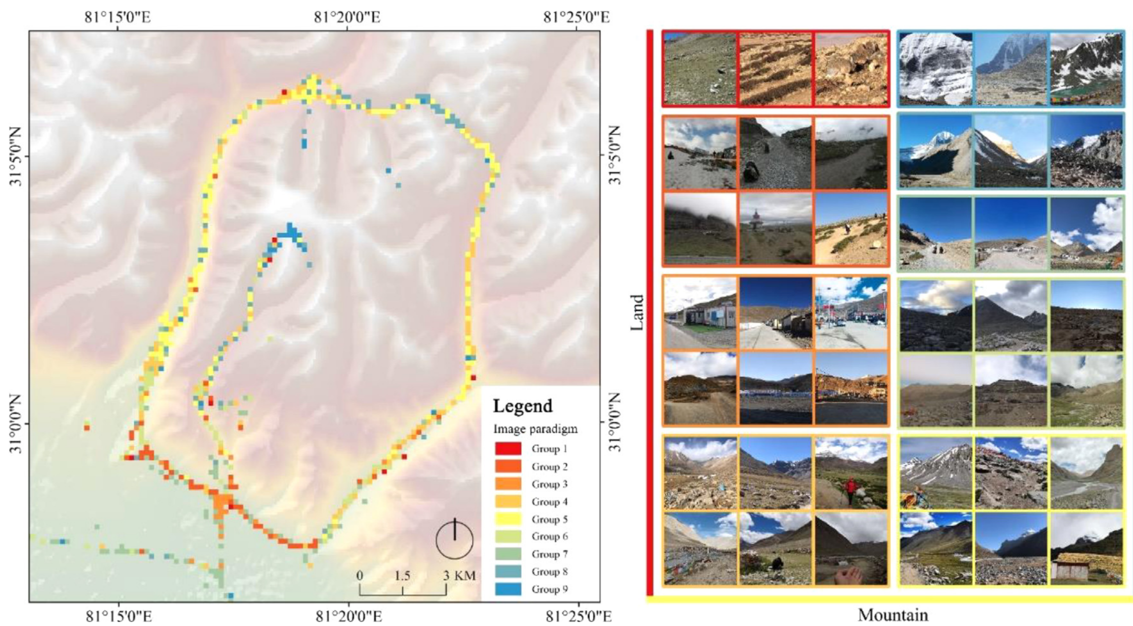

Fig. 14. The distribution of paradigms and typical photos of each.

thority and so on. Mainly taking a stand with tourists and believers in a web-based context, this rudimentary research hopes to partly reveal how Kailash stands out from the landscape and is represented through filtered photos. Combined with the spatial evolution law of perception and behavior of Kailash Sacred Landscape, this paper can provide reference for landscape development and computational methods application in the following aspects.

4.1. Application of social media in landscape image building and cognition

Kailash Sacred Landscape is a valuable historical and cultural heritage for all human beings. Its visual image shapes people's cultural identity and sense of belongings in the form of intangible assets. Using social media to publicize the image of Kailash Sacred Landscape can be a feasible and efficient way. On the premise of maintaining the historical and cultural characteristics of the landscape, the image diversity of Kailash Sacred Landscape should also be enriched. On the one hand, in order to guide the transmission of social media photos, it is necessary to develop some distinctive and representative paradigms and eliminate images that are inconsistent with the landscape character. On the other hand, it is also necessary to open for multi-expression of landscape image so as to avoid only stereotypical image. By analyzing the inherent paradigms of Kailash through social media, it will help to promote the development of visual image in a both coordinated and diversified direction.

Meanwhile, as a new and powerful data source, social media provides a real-time perspective for landscape image cognition. Besides the visual content information, it also contains many additional metadata such as users' background, space-time trajectory and evaluation information. Therefore, it can provide a more diversified perspective compared to traditional data types in many aspects like identifying and interpreting landscape image characteristics, aesthetic value, tourists' space-time behavior, etc. This study uses social media photos to recognize the visual image of Kailash Sacred Landscape, which also proved the feasibility of applying social media in landscape image cognition. By making full use of the advantages of real-time and massive social media, we can explore a feasible way for the digital management of cultural heritage, providing people's real-time cognition and feedback on the landscape.

4.2. Application of the computational methods in landscape research

Under the large background of a shift towards computational methods application in the digital humanities (Hall, 2013), this research also explored and extended the way on applying GIS and AI technology into cultural landscape research. The application of newly developed methods sets this research apart from the previous studies. The AI technology facilitates the data analyzing of huge amounts of photos in a time-saving and standardized way, which provides great flexibility for visual content analysis. Computer vision technology including semantic segmentation not only enables us to recognize the 150 common items from photos, but also the specific proportion of elements which can hardly be calculated manually. Combined with clustering analysis, a new trail is blazed for extracting paradigms in large amounts of photos.

However, shortcomings exist in the application of computational methods in this research as well. The 150 categories of items are still too rough to identify those with religious elements like "sutra streamers" and "prayer stone" from the photos, little in their amounts but significant in the meanings. Besides, content analysis mainly based on the proportion of visual elements cannot distinguish the background and subject in the photo so clearly, which partly lose sight of the meaning of a photo. Finally, lack of the consideration of other factors like color and composition of a photo, the identification of paradigms through clustering is still a rather generalized one, and some more symbolized images of Kailash still need to be filtered out in the further study.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yucheng Zhang: Conceptualization, Resources, Methodology, Software, Investigation, Visualization, Writing – original draft. **Jie He:** Funding acquisition, Supervision, Validation, Writing – review & editing.

Acknowledgements

This research is supported by National Natural Science Foundation of China (project No. 51978448) and Programme of Introducing Talents of Discipline to Universities of Ministry of Education of China (project No. B13011).

References

Balomenou, N., Garrod, B., 2019. Photographs in tourism research: prejudice, power, performance and participant-generated images. Tour. Manag 70201–70217. doi:10.1016/j.tourman.2018.08.014.

Chen, X.Q., Zhang, H.H., 2015. The use of social media in tourism: a literature review. Tour. Trib. 30, 35–43. doi:10.3969/j.issn/1002-5006.2015.08.004.

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y (Eds.), ECCV 2018: Computer Vision – ECCV 2018. Springer, Cham, pp. 833–851. doi:10.1007/978-3-030-01234-2_49.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B., 2016. The cityscapes dataset for semantic urban scene understanding. In: Proceedings of the IEEE Conference on Computer Vision And Pattern Recognition (CVPR). IEEE, pp. 3213–3223. doi:10.1109/CVPR.2016.350.

Cosgrove, D., 2003. Landscape and the European sense of sight-eyeing nature. In: Anderson, K., Domosh, M., Thrift, N., Pile, S. (Eds.), Handbook of Cultural Geography. SAGE Publications, London, pp. 249–268.

Dai, G.Q., Chen, X., 2010. The visual representation analysis of Shangxiajiu street in Guangzhou city: from the perspective of the photos on Internet for marketing. Hum. Geogr. 5, 148-153+91. doi:10.13959/j.issn.1003-2398.2010.05.030.

Garrod, B., 2009. Understanding the relationship between tourism destination imagery and tourist photography. J. Travel Res. 47, 346–358. doi:10.1177/0047287508322785.

Hall, G., 2013. Towards a postdigital humanities: cultural analytics and the computational turn to data-driven scholarship. Am. Lit. 85, 781–809. doi:10.1215/00029831-2367337.

Jiang, K., Yin, H., Wang, P., Yu, N., 2013. Learning from contextual information of geotagged web photos to rank personalized tourism attractions. Neurocomputing 119, 17–25. doi:10.1016/j.neucom.2012.02.049.

Liu, L., Wang, H., 2019. Application of computer vision in urban studies: review and prospect. City Plan. Rev. 43, 117–124.

Llobera, M., 2003. Extending GIS-based visual analysis: the concept of visualscapes. Int. J. Geogr. Inf. Sci. 17, 25–48. doi:10.1080/713811741.

Lo, I.S., McKercher, B., 2015. Ideal image in process: online tourist photography and impression management. Ann. Tour. Res. 52, 104–116. doi:10.1016/j.annals.2015.02.019.

Manovich, L., Douglass, J., 2011. Visualizing change: computer graphics as a research method. In: Grau, O. (Ed.), Imagery in the 21st Century. MIT Press, Cambridge, Mass, pp. 315–338. doi:10.7551/mitpress/9780262015721.003.0017.

McKay, A., 2015. Kailas Histories: Renunciate Traditions and the Construction of Himalayan Sacred Geography. Brill, Leiden.

National Catalogue Service for Geographic Information, 2021. 1: 250,000 National Basic Geographic Database. https://www.webmap.cn/main.do?method=index/, (accessed 17 March 2021).

Olwig, K.R., 1996. Environmental history and the construction of nature and landscape: the case of the 'Landscaping' of the Jutland heath. Environ. Hist. 2, 15–38. doi:10.3197/096734096779522464.

PixelLib, 2021. Semantic segmentation of images with PixelLib using Ade20k model. https://pixellib.readthedocs.io/en/latest/image_ade20k.html/, (accessed 11 March 2021).

Richards, D.R., Friess, D.A., 2015. A rapid indicator of cultural ecosystem service usage at a fine spatial scale: content analysis of social media photographs. Ecol. Indic. 53, 187–195. doi:10.1016/j.ecolind.2015.01.034.

Richards, D.R., Tunçer, B., 2018. Using image recognition to automate assessment of cultural ecosystem services from social media photographs. Ecosyst. Serv. 31, 318–325. doi:10.1016/j.ecoser.2017.09.004.

Rose, G., 2016. Visual Methodologies: an Introduction to Researching with Visual Materials, 4th ed. SAGE Publications, London.

Starrs, P.F., 1997. The sacred, the regional, and the digital. Geogr. Rev. 87, 193–218. doi:10.2307/216005.

Stepchenkova, S., Zhan, F., 2013. Visual destination images of Peru: comparative content analysis of DMO and user-generated photography. Tour. Manag. 36, 590–601. doi:10.1016/j.tourman.2012.08.006.

The SciPy community, 2021. Hierarchical clustering (scipy.cluster.hierarchy). https://docs.scipy.org/doc/scipy/reference/cluster.hierarchy.html#module-scipy.cluster.hierarchy/, (accessed 18 March 2021).

Thornes, J.E., 2004. The visual turn and geography (response to Rose 2003 intervention). Antipode 36, 787–794. doi:10.1111/j.1467-8330.2004.00452.x.

Tobler, W., 1993. Three presentations on geographical analysis and modeling: non-isotropic modeling. Speculations on the Geometry of Geography; and Global Spatial Analysis, (93-1). National Center for Geographic Information and Analysis, Santa Barbara.

UNESCO World Heritage Centre, 2017. Scenic and Historic Area of Sacred Mountains and Lakes http://whc.unesco.org/en/tentativelists/6187/.

Urry, J., Larsen, J., 2011. The Tourist Gaze 3.0. SAGE Publications, London.

Van Gorp, B., Béneker, T., 2007. Holland as other place and other time: alterity in projected tourist images of the Netherlands. GeoJournal 68, 293–305. doi:10.1007/s10708-007-9085-9.

Wang, M., Jiang, R.H., Zhu, H., 2017. A review and revelation of the study of visual in human geography. Hum. Geogr. 32, 10–19. doi:10.13959/j.issn/1003-2398.2017.03.002.

Zhang, K., Chen, Y., Li, C.L., 2019. Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: the case of Beijing. Tour. Manag. 75, 595–608. doi:10.1016/j.tourman.2019.07.002.

Zomer, R., Oli, K.P., 2011. Kailash Sacred Landscape Conservation Initiative: Feasibility Assessment Report. The International Centre for Integrated Mountain Development (ICIMOD), Kathmandu, pp. 7–12.