

DISCRIMINATION OF CULTURAL RELICS SIMILARITY BASED ON PHASH ALGORITHM AND SIFT OPERATOR

Wei Hua ¹, Miaole Hou ^{2,3,4}*, Su Yang ⁵ Youqiang Dong ^{2,3,4}

¹ China University of Mining & Technology, Beijing, 100083 - huawei199299@qq.com

² School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, 102616, Daxing District, Beijing, - (houmiaole, dongyouqiang)@bucea.edu.cn

³ Beijing Key Laboratory For Architectural Heritage Fine Reconstruction & Health Monitoring, Daxing District, Beijing, 102616

⁴ Engineering Research Center of Representative Building and Architectural Heritage Database, Ministry of Education, 102616, Daxing District, Beijing

⁵ North China Institute of Computing Technology, 100083, Beijing, China, - 10770317@qq.com

Commission II, WG II/8

KEY WORDS: pHash Algorithm, Sift Operator, Cultural Relic Similarity, Image Recognition, Cultural Relic Protection

ABSTRACT:

Among the immovable cultural relics such as grottoes, stone carvings of the same period usually show similar artistic styles and repetitive statues. Through the study of the similarity between cultural relics in the same period, we can have reference value for the historical significance, artistic characteristics and cultural relics restoration of cultural relics. The traditional discrimination of cultural relics similarity often depends on the subjective experience judgment of experts, which is time-consuming and inefficient. Often, different experts may issue different conclusions, the lack of convincing objective evaluation basis, can not provide effective scientific support for the protection and restoration of cultural relics. In order to provide objective criteria for evaluating the similarity of cultural relics and avoid the interference of human factors, this paper proposes the concept of similarity index of cultural relics based on perceptual Hash (pHash) algorithm. For orthophoto images of cultural relics with similarity index within the threshold, the SIFT operator is further used to find the corresponding feature points. This paper takes the Buddha statues in Yungang Grottoes as the research object, objectively calculates the similarity index of cultural relics, combines the pHash algorithm and Sift operator, realizes the objective quantitative discrimination of cultural relics similarity, effectively avoids the subjective influence of experts on the judgment of cultural relics similarity.

1. INTRODUCTION

Yungang Grottoes are located on the south side of Wuzhou Mountain, 17 kilometers west of Datong City, Shanxi Province, northern China. There are 45 main caves and more than 51000 stone Buddha statues, which is one of the largest ancient grottoes in China and one of the four major grottoes in China. The statues of Yungang Grottoes are magnificent and rich in content. they can be called the crown of Chinese stone carving art in the 5th century AD and are known as the treasure trove of ancient Chinese sculpture art. Cave 18 of Yungang Grottoes is the largest, most complex and most artistic cave in Yungang Grottoes, and its research significance is self-evident. The main Buddha statue on the northern wall of Cave 18 symbolizes the Emperor Taiwu of the Northern Wei Dynasty. There are many almost identical Buddhas carved on his robes (Fig. 1), in memory of the many monks who lost their lives in the event of extermination of Buddha by Emperor Taiwu.

By judging the similarity between these statues, we can not only find the historical background, political significance, carving technology and other deep meanings of the period in which they lived, but also have scientific guiding significance for the restoration of cultural relics. Under the same environmental conditions, similar cultural relics are often restored with the same technology and materials. At present, experts in the field of cultural relics are often used to judge

the similarity of cultural relics, which is subjective based on their own experience and documentary records. Different experts often have different criteria to distinguish, which leads to the result not having objectivity. In recent years, with the continuous progress of digital photogrammetry and three-dimensional laser scanning technology, the digitization of cultural relics has also made considerable progress. The digitization of cultural relics realizes the permanent preservation of cultural relics archives, and the research on the similarity of cultural relics is more objective and accurate.



Fig. 1 The main Buddha statue on the northern wall of Cave 18

* Corresponding author. E-mail addresses: houlmiaole@bucea.edu.cn (Miaole Hou)

At present, the research on similarity of cultural relics mainly focuses on clustering analysis of cultural relics with similar and significant characteristics. The team of Tokyo University of Japan used cluster analysis method to identify the type of 3D point clouds of facial sculptures from the face of Bayon Temple in Cambodia (KAMAKURA et al., 2008). They explore a more objective way to classify the faces by using measured 3D geometrical models. After alignment of 3D faces in the same coordinate system, orientation, and normalization, they captured in-depth images of each face and then classified them by several statistics methods. However, the acquisition of three-dimensional point clouds requires professional instruments, which are expensive and complex to operate. Zhou Mingquan's team of Northwest University clustered the Terracotta Warriors fragments (Zhang et al., 2018), extracted the contours and surface features of the fragments, so as to facilitate the virtual splicing of the fragments. However, the focus of the team's research is to classify the terracotta warriors and horses fragments according to pre-established templates, without involving the similarity discrimination between single fragments.

In order to solve the problem that there is no objective quantitative index to judge the similarity of cultural relics at present, this paper takes the small Buddha statue on the main Buddha statue of the northern wall of Yungang Grottoes as the research object, and uses the Orthophoto Image of the small Buddha statue as the data source, which has the characteristics of non-contact, low cost and easy operation. Data acquisition can be completed only by using the consumption-grade digital camera. In data processing, the complexity of processing is lower than that of using three-dimensional model discrimination. This paper also puts forward the similarity index of cultural relics, and summarizes an automatic identification method of similar Buddha statues combined with pHash algorithm and Sift algorithm.

The outline of the rest of this paper is as follows. Section 2 introduces the general principles of pHash algorithm and SIFT algorithm; In Section 3 we takes the small Buddha carved on the principal Buddha statue Thousand Buddha's Gown in Grotto 18 of Yungang Grottoes as an example, and present experimental results. Finally, in Section 4 we summarize and conclude this paper.

2. A METHOD FOR CALCULATING THE SIMILARITY OF CULTURAL RELICS

2.1 Similarity Index of Cultural Relics Based on pHash Algorithms

The pHash algorithm is a method of forming a set of fingerprint data after image processing. It is often used in image search of finding pictures by pictures. The comparative image with similar contours can be found by referring to the reference image. When searching for pictures, pHash algorithm extracts the features of the input contrast image, generates a set of hash fingerprint data expressed by two-dimensional array, and then compares the fingerprint data with the hash fingerprint data of the target image to find the corresponding image. Compared with other forms of eigenvalues, the two-dimensional data has a higher timeliness (Wang et al., 2017). The pHash algorithm is not affected by the image resolution and color, and has a higher robustness. The processing steps of the pHash algorithm are shown in Figure 2.

(1) Reduce the size of the picture.

The pHash algorithm reduces the size of the image to $N \times N$ (N takes 32 in general), and only retains the basic information such as the structure and shade of the image. The purpose of this method is to simplify the calculation of Discrete Cosine Transform (DCT) by avoiding the difference between image size and image proportion.

(2) Simplifying color.

Transform the image into gray image, further simplify the calculation.

(3) Calculate the average gray level.

The average gray value of all the pixels in the simplified gray image is calculated.

(4) Calculate the average value of DCT.

(5) Calculate the hash value.

This step will generate hash fingerprints and is the core step of the whole algorithm. In the 32×32 matrix generated by DCT transformation, most of the information is located in the 8×8 matrix in the upper left corner. Therefore, only the elements in the 8×8 matrix need to be compared with the DCT mean value. The greater or equal mean value is marked as 1 and the smaller mean value is marked as 0. Thus, a 64-bit hash fingerprint composed of 0 and 1 is generated.

In the above algorithm, two-dimensional discrete cosine transform (DCT) is derived from the coefficient sub-matrix of 8×8 in the matrix.

$$G_{u,v} = \frac{1}{4} \alpha(u)(v) \sum_{x=0}^N \sum_{y=0}^N g_{xy} \cdot \cos\left[\frac{(2x+1)u\pi}{N^2}\right] \cos\left[\frac{(2y+1)v\pi}{N^2}\right] \quad (1)$$

Where:

G = the $N \times N$ image pixel;

g = the threshold matrix of $N \times N$ matrix.

α = a cosine coefficient matrix.

If it represents the matrix form, formula (1) can be simplified as follows:

$$F(u,v) = A f(x,y) A^T \quad (2)$$

After getting 64-bit hash fingerprint, we need to compare it with the hash fingerprint of the target image. In this paper, we propose the concept of similarity index of cultural relics. The similarity index of cultural relics is to calculate the number of different characters corresponding to the 64-bit hash fingerprint of two images by calculating the Hamming distance. When the similarity index of cultural relics is less than 5, it is judged that the cultural relics represented by the two images have similar relationship, and the smaller the value, the greater the similarity; when the similarity index of cultural relics is greater than 10, it is judged that the cultural relics represented by the two images belong to totally different cultural relics.

2.2 Feature point matching based on SIFT algorithm

Similarity index of cultural relics can only be used as a preliminary screening result of whether cultural relics are similar or not. The disadvantage is that it does not accurately find out the similarity feature points between two cultural relics images. Therefore, SIFT operator is needed to extract the similarity feature points from the two cultural relics images which are preliminarily screened as similar cultural relics.

SIFT (Scale-invariant feature transform) was first published by David Lowe at the International Conference on Computer Vision in 1999 and improved in 2004 (Lowe, 2004). SIFT operator has the invariance of image rotation and scale

transformation, and is also robust to affine transformation, brightness transformation and noise (Gu et al., 2015). It is often used to match the feature points between pictures, so it can extract the detailed feature points of Buddha's face, realize the objective evaluation of Buddha's similarity, and provide a scientific basis for the restoration of cultural relics. SIFT operator detects the feature points of the input image, generates the set of feature points, and then matches the set of feature points of different images to get similar feature points. The core content of SIFT operator is the construction and detection of feature points. Its main steps are as follows:
(1) Constructing Scale Space
The scale space obtained by using the convolution of Gauss function is as follows:

$$L(x, y, k\delta) = G(x, y, k\delta) * I(x, y) \quad k = 1, 2 \dots S \quad (3)$$

Where:

$I(x, y)$ represents the coordinates of the image I space, $k\delta$ is the scale factor. The magnitude of δ is related to the smoothness of the image. The smaller the value of δ , the finer the scale, and the higher the resolution.

Then the DOG Scale-Space is constructed:

$$D(x, y, \delta) = (G(x, y, k\delta) - G(x, y, \delta)) * I(x, y) \\ = L(x, y, k\delta) - L(x, y)$$

(4)

(2) Finding Extremum Points in Scale Space

Comparing each sampling point with all 26 adjacent points in the DOG space, the point is considered as a feature point in the current scale when it is the maximum or minimum value.

(3) Accurate positioning of extremum points

Using Harris Corner detector (Harris and Stephens, 1988), the spatial scaling function is expanded by Taylor expansion:

$$D(x, y, \delta) = D(x, y, \delta) + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (5)$$

The exact position can be obtained by calculating the derivative of formula (5) and making it 0:

$$\hat{x} = -\frac{\partial D^{-1} \partial D}{\partial x^2 \partial x} \quad (6)$$

Bring formula (6) tantulum into formula (5) and take only the first two:

$$D(\hat{x}) = D(x, y, \delta) + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (7)$$

When $|P(\hat{x})| \geq 0.03$, the feature is retained or discarded. Then the ratio of principal curvature is checked, and the feature points less than the threshold are retained. The default threshold is adopted in this paper.

(4) Constructing Directional Parameters of Characteristic Points

The modulus and direction formulas of the gradient at (x, y) are:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \\ \theta(x, y) = \arctan 2 \sqrt{(L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))} \quad (8)$$

From this, we can get three key information of feature points: location, scale and direction, and finally generate an $n \times 128$ -dimensional SIFT description vector, where n is the number of feature points.

3. EXPERIMENTS

Taking the main Buddha statue on the northern wall of Cave 18 of Yungang Grottoes as an example, this paper selected the statue of Buddha carved on its body, which was covered with a thousand Buddha robe. The experimental procedure is shown in Fig. 3. The original image data were collected with Nikon D810 digital camera, and the Orthophoto Image was generated. The main flow chart is as follows (Fig. 2).

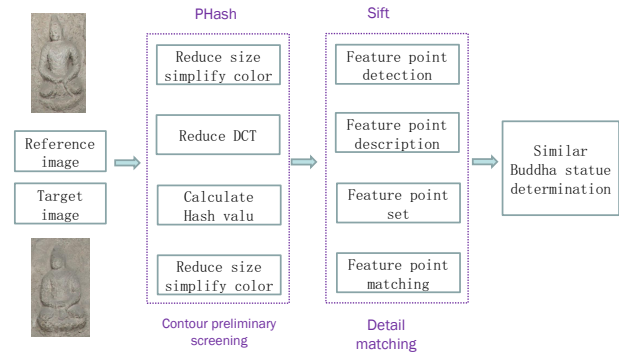


Figure. 2 flow chart of similarity discrimination algorithm

The first step is image acquisition and orthophoto generation. Among the many small Buddha statues, the well-preserved ones are selected as reference images and the other four groups of small Buddha statues are used as contrast images (Fig. 3).



(a) The position of Buddha statues on the robes of the Great Buddha



(b) Detailed picture of Buddha statue

Fig. 3 Similar Buddha statues on the cassock of the great Buddha in the grottoes

The second step is to pre-process the selected orthophoto images of small Buddha statues. The orthophoto images of single Buddha statues that need to be compared are cut out to avoid affecting the accuracy of subsequent experiments (Fig. 4).

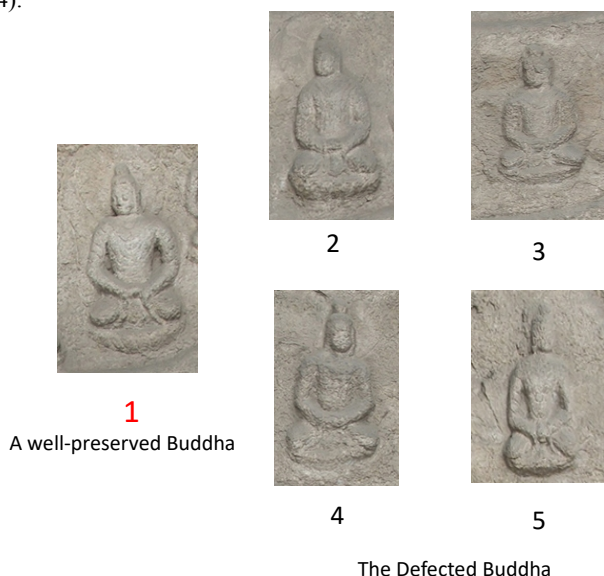


Fig. 4 Contrast Buddha statues and reference Buddha statues

The third step is to use the pHash algorithm to compare the small Buddha image of the contrast group with the reference image one by one, and to find the similarity index of cultural relics between each group, retain the image with the similarity index of less than 5, and enter the next step (Fig. 5). According to the calculation results, the similarity index of the cultural relics in the comparison group 2-1 is 3; the similarity index of the cultural relics in the comparison group 3-1 is 3; the similarity index of the cultural relics in the comparison group 4-1 is 2; the similarity index of the cultural relics in the comparison group 5-1 is 1.

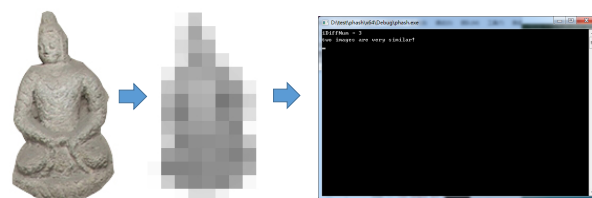
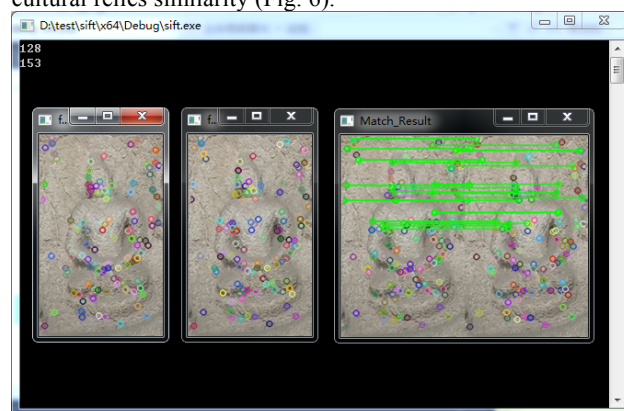


Fig. 5 Computation of Cultural Relics Similarity Index

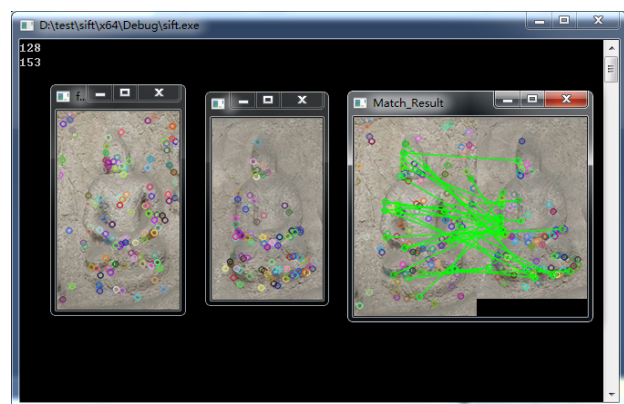
The fourth step is to implement SIFT algorithm for the image of Buddha whose similarity index is less than 5. Similar feature points of two images can be obtained directly by SIFT algorithm, so as to assist cultural relics experts in identifying cultural relics similarity (Fig. 6).



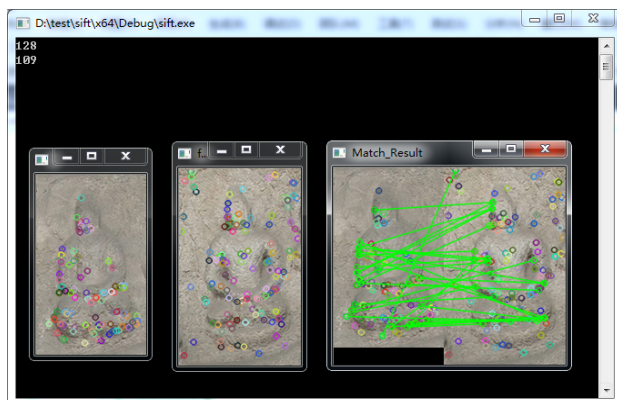
(a) Matching of Characteristic Points between Contrast Buddha Statue 2 and Reference Buddha Statue 1



(b) Matching of Characteristic Points between Contrast Buddha Statue 3 and Reference Buddha Statue 1



(c) Matching of Characteristic Points between Contrast Buddha Statue 4 and Reference Buddha Statue 1



(d) Matching of Characteristic Points between Contrast Buddha Statue 5 and Reference Buddha Statue 1

Figure 6 Extraction of similar Buddha statue detail feature points using Sift algorithm

4. CONCLUSIONS

This paper takes the small Buddha statue on the thousand Buddha robe on the north wall of the 18th cave of Yungang Grottoes as the research object, and puts forward an objective method to determine the similarity of cultural relics. It solves the problem that the similarity judgment of cultural relics depends on the subjective judgment of experts, but there is no objective quantitative index. In the research, we summarize an automatic identification method of similar Buddha statues which combines the pHash algorithm with the Sift algorithm. By using the pHash algorithm, the unique hash fingerprint can be generated quickly and accurately. Similar cultural relic's images can be screened preliminarily. The similarity index of cultural relics based on the Hamming distance method is proposed, and the similarity degree of cultural relics is objectively and quantitatively given. Using the scale invariant property of SIFT operator, this paper also gives the specific similarity feature points for the preliminary screening of similar cultural relics images for the reference of cultural relics experts.

REFERENCES

- Gal, R. , & Cohen-Or, D., 2006. Salient geometric features for partial shape matching and similarity. *ACM Transactions on Graphics*, 25(1), 130-150.
- Gu Mingtao, Zhang Dinghui, and Liu Jing., 2015. Image Sensing Hash Algorithms Based on Cluster Analysis and SIFT. *Data Communication* 3.
- Harris, C. G., & Stephens, M. J., 1988. A Combined Corner and Edge Detector. *Proc of Fourth Alvey Vision Conference*.
- Kamakura, M., Oishi, T., Takamatsu, J., & Ikeuchi, K., 2005, October. Classification of Bayon faces using 3D models. In *Virtual Systems and Multimedia* (pp. 751-760).
- Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.
- Lu, M., Kamakura, M., Zheng, B., Takamatsu, J., Nishino, K., & Ikeuchi, K. 2011, October. Clustering Bayon face towers using restored 3D shape models. In *2011 Second International Conference on Culture and Computing* (pp. 39-44). *IEEE*.
- Lu, M. , Zheng, B. , Takamatsu, J. , Nishino, K. , & Ikeuchi, K., 2011. Preserving the khmer smile: classifying and restoring the faces of bayon.
- Monga, V. , & Evans, B. L., 2004. Robust Perceptual Image Hashing Using Feature Points. *International Conference on Image Processing. IEEE*.
- Wang Tuo, Yu Xuhong, and Liu Zhijie., 2017. Research on Hash Fast Image Matching Algorithms. *Journal of Chongqing University of Science and Technology (Natural Science Edition)* 3.
- Yang, B., Gu, F., & Niu, X., 2006, December. Block mean value based image perceptual hashing. In *2006 International Conference on Intelligent Information Hiding and Multimedia* (pp. 167-172). *IEEE*.
- Zhang, Y., Li, K., Chen, X., Zhang, S., & Geng, G., 2018. A multi feature fusion method for reassembly of 3D cultural heritage artifacts. *Journal of Cultural Heritage*, 33, 191-200.