



Available online at
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com/en



Original article

Average-face-based virtual inpainting for severely damaged statues of Dazu Rock Carvings



Haiyan Wang^{a,*}, Zhongshi He^b, Yiman He^c, Dingding Chen^b, Yongwen Huang^d

^a Department of Art History, Sichuan Fine Arts Institute, 401331 Chongqing, PR China

^b College of Computer Science, Chongqing University, 400044 Chongqing, PR China

^c Faculty of Architecture and Urban Planning, Chongqing University, 400044 Chongqing, PR China

^d School of Electrical and Information Engineering, Chongqing University of Science and Technology, 401331 Chongqing, PR China

ARTICLE INFO

Article history:

Received 21 February 2018

Accepted 6 August 2018

Available online 28 August 2018

Keywords:

Average face

Symmetry

Image deblurring

Exemplar-based inpainting

Dazu rock carvings

ABSTRACT

Numerous image inpainting algorithms are guided by a basic assumption that the known region in the original image itself can provide sufficient prior information for the guess recovery of the unknown part, which is not often the case in actual art image inpainting. In order to solve the challenging inpainting case that there is little image prior in the remainder of the original image, we propose an average-face-based inpainting method based on a sample database with 3 steps: reference images selection, average image generation and exemplar-based image inpainting. In which, average image generation is crucial. In the inpainting framework, the average image can be directly viewed as an inpainting proposal for the severely damaged or absolutely lost image. Moreover, the average image can be applied to exemplar-based inpainting algorithm as a sample image to extend the searching region for match patch, so as to perform the restoration for images with large-scale or irregularly damaged holes. The inpainting experiments over some facial images of Dazu Rock Carvings demonstrate the validity and effectiveness of our method. It is first utilized for two extremely challenging inpainting tasks: reconstruction for the stolen head of Willow Avalokitesvara in Shimenshan No. 6 and the absolutely broken heads of two Avalokitesvaras in Beishan No. 180. Compared to the failure of the exemplar-based inpainting algorithm within the original image and the directly duplication of a similar image, the generated average image can be as a more reliable inpainting proposal. The comparative experiments also show the efficiency and advantage of the average face applied to exemplar-based inpainting framework. Compared with some related inpainting algorithms, our method is more competitive when there is little prior information in the original image. The efficient virtual inpainting results are valuable references for both cave art historians and conservators.

© 2018 Elsevier Masson SAS. All rights reserved.

1. Introduction and motivation

As the pinnacle of Chinese cave temple art dating from the 9th to the 13th century, Dazu Rock Carvings in Chongqing is well known as the cultural treasure of carvings in Tang and Song dynasties. However, due to moist climate, shallow grotto and niche depth, the carvings have been suffering from natural erosion for almost one thousand years, so the inpainting is urgently needed. In order to reduce the risk of the physical inpainting and keep the carvings' original appearance, it is necessary to introduce digital image processing techniques to perform virtual inpainting initially.

Numerous image inpainting algorithms have been proposed since Bertalmio et al. [1] first introduced art restoration terminology "image inpainting" to computer science in 2000. It refers to a guess recovery of the missing area in an image according to the prior information extracted from its known region. According to different image priors, the inpainting approaches can be classified as 3 different categories: diffusion-based inpainting methods [2–4] focusing on parameter models to propagate the structural components of an image via smoothness priors; exemplar-based inpainting algorithms [5–9] compositing match patches to reconstruct the unknown region according to self-similarity priors; and sparsity-based inpainting methods [10–12] relying on sparsity priors. Experimental results show that the diffusion-based and sparsity-based methods work well for the recovery of structures, small and sparsely distributed gaps. Whereas exemplar-based

* Corresponding author.

E-mail address: why@scfai.edu.cn (H. Wang).

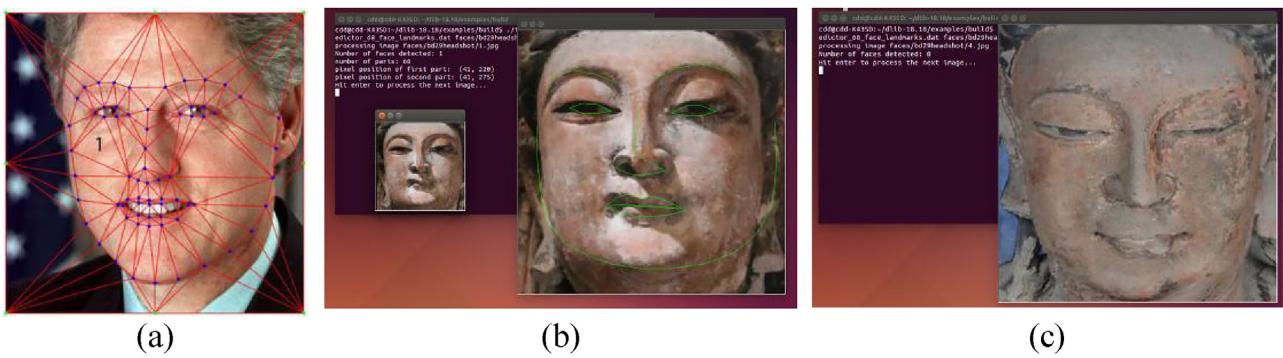


Fig. 1. Facial detection via Dlib.

methods work efficiently in the restoration of large-scale texture region with homogeneous or regular patterns. Therefore, the latter is widely used in many types of artwork restoration, such as paintings [13–15], frescoes [16–18], aging photographs [19] and film sequences [20,21]. Virtual restoration can be valuable to art historians for getting more information, and useful to art conservators for its guidance to the physical restoration.

However, almost all the existing inpainting approaches are enforced in a single image, guided by a basic assumption that the remainder of the original image itself can provide sufficient information for the guess recovery, which is not often the case for art image inpainting. In some cases, the art image itself is so badly damaged that there is little prior information as a good model for the inferring, or the original image is with large-scale or irregularly dispersed holes, some damaged statues of Dazu Rock Carvings for example.

Archaeological and art history scholars reach an agreement that Buddhist images made in a specific period and location share some universal characteristics. It is quite accord with the idea of “average face” which under the hypothesis that certain groups of people may have common facial characteristics. Song Li, the distinguished Buddhist art researcher in China, sorted out the main clues of the development of Bodhisattva images in China [22]. He pointed out that the model style of Bodhisattva statues evolved with the spread and development of Buddhism. Yuhua Lei, the director of Buddhist Archaeology Institute of Chengdu Cultural Relics and Archaeology Research Institute, made an overview [23] and concluded unique features of Buddhist images in different locations of Sichuan from North and South Dynasty to Ming and Qing Dynasties. Most Dazu Rock Carvings were made in Southern Song Dynasty and had the unique esoteric Buddhism style. Furthermore, Buddha statues cannot be built at will and there are detailed descriptions for Buddhist imagery in Buddhist scriptures. The strict metrics and measurements of Buddhist image-making in China were concluded in Buddha’s Theory of Imagery in Qing Dynasty. Moreover, our valuable findings in [24] also show that the statues in the same cave or on the same subject in Dazu are similar.

At the same time, an interesting finding in evolutionary biology is that both adults and infants have the cognitive ability of averaging the individual exemplars within a category to create a prototype. Accordingly, experienced art image restorers can utilize the ability to guide their inpainting task when there is little prior information in the original image.

Therefore, inspired by the seminal work of Hays and Efros [25] and other researches [26–28] which performed inpainting tasks via multiple Internet images rather than a single original image, we propose an average-face-based inpainting method based on a sample database to solve the aforementioned challenging inpainting cases for facial image of Dazu Buddhist statues.

2. Related work

2.1. Average face

The term “average face” is a representative or a prototype of a category of faces. It is mainly used to discuss physical attractiveness in cognitive and developmental psychology and evolutionary biology. Its initial intention is to search for the answer of what constitutes beauty, or what defines a beautiful face.

The pioneer study of averageness was Galton (1878, 1883) and Stoddard (1886, 1887). They composed portraits by superimposing photographic exposures of faces. Despite of Galton’s initial aim of creating graphic representations of types of faces, such as criminals, meat-eaters, vegetarians, and tuberculosis patients, they observed that the composites were more attractive than the individual components. Their observations were forgotten for over a century until 1990. Using computerized technique of mathematically averaging faces over the photographic, psychologists Langlois and Roggman [29] systematically examined whether the computer-average-generated composite image is linked with facial attractiveness. They just simply averaged the pixel matrices of the individual faces to create prototypes.

Based on recent success of the face detection technology, Mallick described the steps for generating an average face as follows:

- step 1: facial feature detection;
- step 2: coordinate transformation;
- step 3: face alignment;
- step 4: face averaging.

For each facial image, 68 facial landmarks are firstly calculated (shown in blue in Fig. 1(a)) via Dlib, a cross-platform software library with a set of tools for dealing with different tasks in a wide range of domains of machine learning.

The input facial images with different sizes are then normalized to be the same size and brought to the same reference frame (two corners of the eyes are aligned). To achieve this, the faces are warped to a 600×600 image using warp affine.

For the alignment of other facial features, 68 landmark points and 8 points on the boundary of the image (shown in green in Fig. 1(a)) are used to divide the images into triangular regions by calculating a Delaunay triangulation.

After above steps applied to all the input images, the average image is calculated by simply adding the pixel intensities of all the warped images and divide by the number of images.

2.2. Image deblurring

The blur process is generally modeled as

$$B = I \otimes k + \varepsilon \quad (1)$$

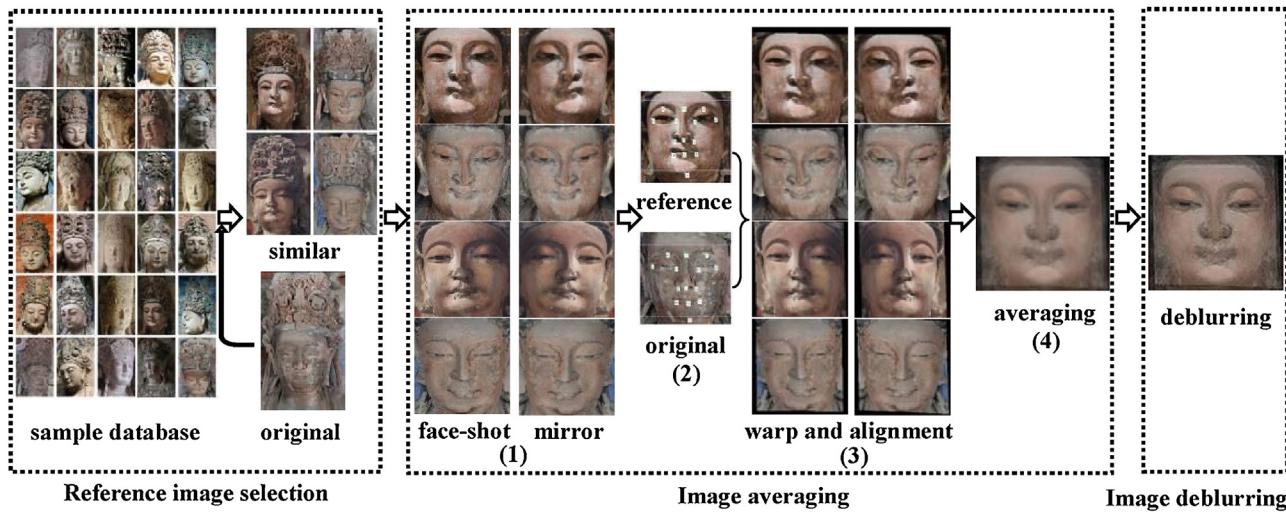


Fig. 2. Framework of the average image generation based on a sample database.

where I is the latent image, k is the blur kernel, a blur point-spread-function (PSF), ε is the image noise, \otimes denotes convolution and B is the blur image. Image deblurring therefore can be viewed as a deconvolution process.

Xu and Jia [30] presented an efficient two-phase sparse kernel estimation for robust motion deblurring in an iterative optimization scheme. They first introduced Gaussian priors to estimate the blur kernel. It is an automatic gradient selection algorithm to exclude the detrimental structures and preserve useful sharp edges in latent image restoration. They then employed an iterative support detection (ISD) algorithm for the kernel refinement, a powerful numerical scheme through iterative support detection, to further improve the result quality with adaptive regularization. For the final deconvolution step, they developed a fast $TV-\ell_1$ deconvolution method to efficiently reject outliers and preserve structures. Thus the latent image is coarsely restored.

3. Average image generation based on a sample database

Built upon average face generation, we propose an average image generation based on a sample database in following stages:

- stage 1: reference image selection;
- stage 2: image averaging;
- stage 3: image deblurring.

The framework is illustrated in Fig. 2. Reference image selection can greatly reduce time complexity of the image reconstruction based on a sample database. As the similar images of the original image (the image needed to be restored) can be efficiently selected via VGGNet-based recognition method [24]. Therefore the key work of this paper is stages 2 and 3.

3.1. Image averaging

Image averaging is the crucial stage in the framework. The pixels of the reference images are averaged to create an image containing the universal traits. For the best visual quality, symmetry, facial key feature landmark, image warp and alignment, and pixel averaging are necessary steps.

3.1.1. Symmetry consideration

Findings from both neuroimaging and behavioral studies suggest that symmetry may contribute to the attractiveness of average

front view-faces. The computer graphic studies [31–35] also demonstrate that symmetry can have a positive influence on attractiveness.

However, if the composite images are with different angles, warping is needed for a coordinate conformity. It inevitably introduces distortion in the average face if the angular deviation is large. Therefore, symmetry should be considered. To ensure the symmetry of the average face, similar images and the corresponding mirror images are together to be reference images. Mirror images can be obtained by flipping horizontally.

Langlois and Roggman statistically found that the 8- to 16-composite face was the most visually attractive of all the faces according to the illustrated rating for attractiveness. Thus, for the number of the similar images, we usually select top 4 to 8 from the database in our experiments except of special case. The reference images are becoming 8–16 together with their mirror images, shown in Fig. 2(1). It is sufficient for an average face to attain both average values and high attractiveness ratings.

It is worthy to mention that a face shot of each selected similar image is made to extract the facial part, aiming to delete unrelated information and normalize them to be the size of 300×300 .

3.1.2. Facial key feature landmark

As mentioned in Sect. 2.1, it is badly constrained by the strict request of angle consistent for a better prototype of a specific category. Thus the most important step is to warp the reference images for alignment, and the feature landmark is crucial to the warping procedure. Studies have demonstrated that even very subtle differences in apparent skin quality can have pronounced effects on facial attractiveness [36,37]. Accordingly, smoothness of skin texture is critical for the facial feature detection in Dlib. Illustrated in Fig. 1, a successful landmark in (b) contributes to the features on the face is quite distinctive for the skin texture is smooth. While it fails to detect the key points in (c) due to the disturbing information on the facial skin. Such facial skin damage is usual for Dazu rock carvings, so it is impossible to directly employ the Dlib to mark the facial features.

Therefore, we have to mark 14 key points of the eyes, nose and mouth manually, illustrated in Fig. 2(2). Experimental results show that it is sufficient for the alignment.

3.1.3. Image warp and alignment

After marking the key features, the original and reference images are warped for alignment using affine transform, illustrated

in Fig. 2(3). The original image is usually as the reference substance if its remainder includes the key features. Otherwise, the reference image with the most similar angle to the original image can be chosen as the reference substance.

3.1.4. Pixel averaging

At last, the average image generated by averaging the pixel values of these aligned reference images with

$$A = \frac{1}{n} \sum_{i=1}^n p_i, n \geq 8 \quad (2)$$

Shown in Fig. 2(4), the average image adopted the universal traits of each reference image is generated. Unfortunately, the blurs in the average image are increasing with the increase of the amount of the involved reference images. Image deblurring is therefore necessary to improve its visual quality and correspondingly the final inpainted result.

3.2. Image deblurring

As the blur in the average image is caused by the angular deviation of the reference images. This kind of blur is quite similar to the motion blur caused by camera shake. The excellent and efficient robust motion deblurring algorithm proposed by Xu and Jia [30] is applied to process the blurs.

Thanks to the authors' volunteer share of an efficient graphic user interface (GUI) online¹, we can easily obtain a deblurring image, shown in Fig. 3(a). For the blur size can be set with different parameter, such as "default", "small", "medium", "large", "custom" and so on. The comparison of the results with different parameters is illustrated in Fig. 3(b). The blur size is empirically set with "small", and with other parameters as "reduce ringing" and "reduce noise" for the best visual quality of the result.

Mirror image adding and image deblurring greatly improve the visual quality of the average image for reducing the distortion and blurs, ensuring its validity and plausibility. The algorithm of the whole work can be sketched in Algorithm 1 with the aforementioned three main stages.

Algorithm 1. Average image generation

Input:

input image I
sample database $D = \{p_1, p_2, \dots, p_n\}$

Output:

feature vector $F = \{f_1, f_2, \dots, f_n\}$
averaged image I_a

1: select similar images $I_r = \{Sim_1, Sim_2, \dots, Sim_n\}, n \in [4, 8]$ of I from D

//Algorithm 1 and formula (2) in [24]

2: flip horizontal I_r and get the mirror images

$I_m = \{Mir_1, Mir_2, \dots, Mir_n\}, n \in [4, 8]$

3: add I_m into I_r : $I_r = I_r \cup I_m$

4: mark the 14 key feature points of I and I_r

5: warp affine and align I_r based on I

6: average the pixels of I_r //formula(2)

7: output I_a

8: deblur I_a

4. Average-face-based inpainting method

As a typical face of a category, an average image can be with the universal facial characteristics of each component image, as well as avoids a rough duplication of any component. Thus, average face can be directly viewed as the reconstructed proposal for an

Table 1

Sample database of Dazu Buddhist statue facial images.

Number of the cave\niche, name and construction time	Total images
Beishan No. 136, Cave of Prayer Wheel, A.D.1142–1146	258
Beishan No. 180, Cave of Thirteen Incarnations of Avalokitesvara, A.D.1116–1122	418
Shimenshan No. 6, Cave of Ten Avalokitesvaras, A.D.1136–1141	608
Baodingshan No. 11, Niche of Sakyamuni Entering Nirvana, A.D.1174–1252	152
Baodingshan No. 18, Sutra of Amitabha and His Pure Land, A.D.1174–1252	402
Baodingshan No. 29, Cave of Full Enlightenment, A.D.1174–1252	374

almost completely damaged image. And it is also can be applied as a source image for match patch in exemplar-based inpainting algorithm when the original image is severely damaged.

After executing exemplar-based inpainting algorithm with the average face, result blending is added to enhance the visual quality. It is because there may be great difference in lighting, contrast and color between the composite match patches and the remainder region of the target image for the source region is an average image based on multiple different images. Blending processing can make the known region and the reconstructed region look more harmonious in vision.

As an excellent gradient domain image processing method, Poisson blending [38] can seamlessly blend an object or texture from a source image into a target image avoiding noticeable difference. It is therefore employed to post-process the restored image.

On the basis of the classical exemplar-based inpainting algorithm proposed by Criminisi in 2004, average-face-based inpainting method based on a sample database for Dazu Buddhist facial image can be described as Algorithm 2.

Algorithm 2 average-face-based inpainting

Input:

target image I
sample database $D = \{p_1, p_2, \dots, p_n\}$

Output:

feature vector $F = \{f_1, f_2, \dots, f_n\}$
averaged image I_a
completed image I_c

1: load target image I and sample database $D = \{p_1, p_2, \dots, p_n\}$

2: execute average image generation, get average image I_{al} // Algorithm 1 in Sect. 3

3: I_a as the source image, execute exemplar-based inpainting// region filling algorithm of Table 1 in [5]

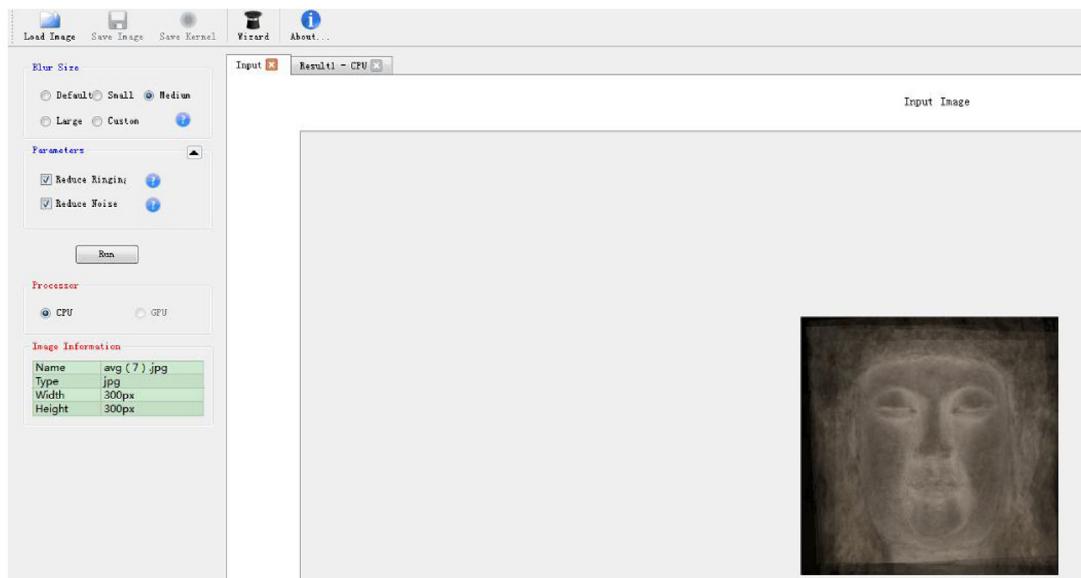
4: blend the restored fragment with the target image I // Poisson blending

5: output feature vector $F = \{f_1, f_2, \dots, f_n\}$, averaged image I_a , and completed image I_c

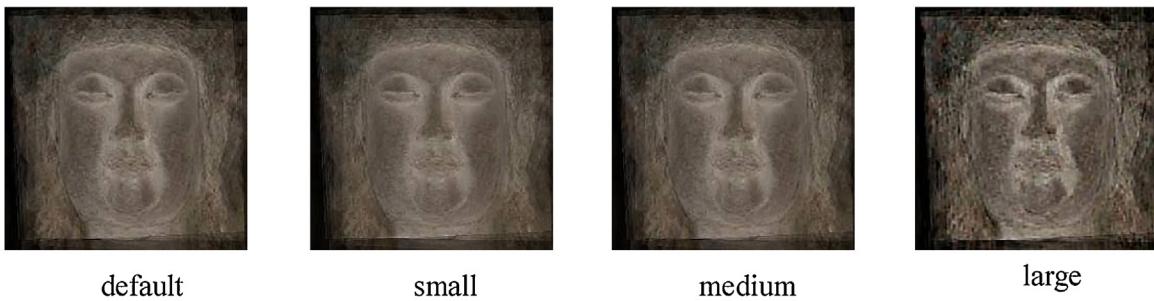
5. Experiments and analysis

In order to validate the effectiveness of average-face-based inpainting method, a variety of experiments over Dazu Buddhist facial image inpainting are enforced. Dazu Rock Carvings are comprised of the cliffside carvings at five regions including Baodingshan, Beishan, Nanshan, Shimenshan and Shizhuanshan. As there is only one statue cave on Taoism at Nanshan, and Shizhuanshan carvings are famous for Confucius subject, the sample database in Table 1 is composed of 2212 Buddhist carvings' RGB head images of 6 different caves or niches at 3 different hills. They are partly provided by the Dazu Institute and mainly photographed by the authors. The head (face and crown) part of each Buddhist statue image is captured in Photoshop with length-width ratio of 5:8. All the algorithms are implemented on Matlab7.8.0 (R2009a).

¹ http://www.cse.cuhk.edu.hk/leojia/projects/robust_deblur/index.html.



(a) GUI of robust deblurring



(b) Different quality after deblurring with different blur size

Fig. 3. GUI of robust deblurring and the result comparison.

We firstly experiment with two challenging inpainting cases that the target image is with no related priors. Some comparative experiments are then set to validate the efficiency of the average image for exemplar-based inpainting.

5.1. Face reconstruction of the Willow Avalokitesvara

As shown in Fig. 4, except of the Three Sanctuary of the West on the front wall, there are ten Avalokitesvara statues on left and right side wall in Shimenshan No. 6, named “the Cave of Ten Avalokitesvaras” accordingly. Unfortunately, the whole head of the Willow Avalokitesvara on the left side wall was stolen, shown in Fig. 4(c).

It fails to restore the head only within the single original image, no matter how classical (denoted as Criminisi) [5] or efficient (denoted as Huang) [39] exemplar-based inpainting algorithms employed, illustrated in Fig. 5(c)–(d). It is also invalid to directly duplicate a face from a similar picture (denoted as Hays) [25], shown in Fig. 5(e), because there is no two statues with the same face in Dazu rock carving.

Luckily, the other 9 Avalokitesvaras in this cave are still survive and preserved well, shown in Fig. 4(b)–(c). Therefore, average-face-based inpainting is introduced. For the loss of the whole head, it is impossible to mark the key features, the front view facial images of the 9 statues are directly chosen as the component images, shown in Fig. 6(a). Then the 9 mirror images are added together as the reference images, illustrated in Fig. 6(b). After 14 key features are marked, one reference image is viewed as the reference substance,

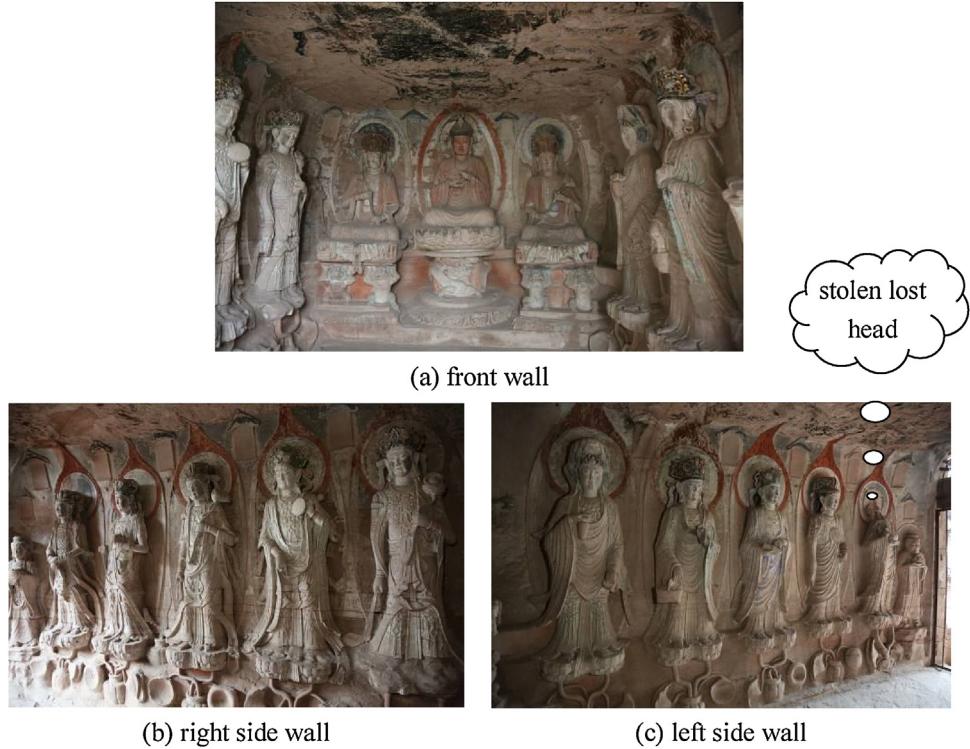
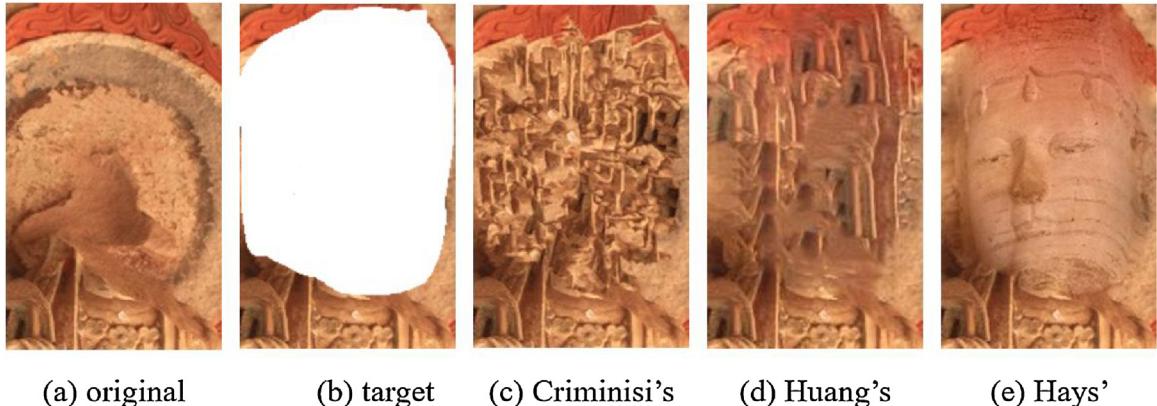
the left 17 reference images are warped for alignment with it, shown as Fig. 6(c). At last, all the aligned references are averaged to be one image in Fig. 6(d). The deblurred average image in Fig. 6(e) can be as a possible face of the Willow Avalokitesvara.

To testify the reliability of viewing the average face of the other 9 Avalokitesvara statues as the face of the Willow Avalokitesvara, three actual Willow Avalokitesvara images with head² (No. 10-1 to No. 10-3 in Fig. 7(a)) are used to make a similarity measurement. The compared average images are generated with 5- and 9-components without mirror images, 12- and 18-components with mirror images, denoted as Avg(5), Avg(9), Avgm(12) and Avgm(18) respectively, shown in Fig. 7(b).

From Fig. 7(a) and (b), we can find that the average images are not the duplication of any component image, but they are similar to the 3 actual image of the Willow Avalokitesvara in visual perception, such as the round face and chin, long and thin eyebrows and eyes, straight nose and slightly raised mouth corner.

Moreover, the Cosine similarity, a measure of similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them, is used to measure the similarity. The smaller the angle is, the value is more approximate 1, and the more similar they are. As shown in Fig. 8, the similarity value highly reaches to 0.858, and the lowest is above 0.814, even though the

² The photos taken in 1990s were provided by Professor Tianxiang Wang of Sichuan Fine Arts Institute.

**Fig. 4.** Cave of Shimenshan No. 6.**Fig. 5.** Failure inpainting result of the Willow Avalokitesvara.

two groups are in different view (the average images are with front view, while the actual images are with left side view).

Furthermore, for the similarity measurement between the average images and the other 9 statues' facial images, the highest value is 0.928 and the lowest is 0.84. Therefore, it is valid and reliable to take the average face of the other 9 statues as the lost face of the Willow Avalokitesvara.

In order to compare the effect of different components on the average face, we make a paired similarity measurement between the average faces with different reference images, such as 5-, 6-, 7-, 8-, 9-component without mirror images, denoted as Avg(5), Avg(6), Avg(7), Avg(8) and Avg(9); 10-, 12-, 14-, 16-, 18-component with mirror images, denoted as Avgm(10), Avgm(12), Avgm(14), Avgm(16) and Avgm(18) in Table 2. The highest (except of 1) and lowest similarity value in each column are highlighted in red bold font and blue Italic font respectively. Their distribution shows that the average faces without mirror components are more similar to each other. Similar results are also revealed for average faces with

mirror components. Moreover, the difference value between the highest and the lowest in each column is in a small range from 0.011 to 0.036. It means that all the average faces are quite similar with each other even though they are generated from different components. It also illustrates that the mirror images play only roles for symmetry consideration. Furthermore, though the similarity values generally decrease with the increase of the involved reference images, it affects little for the quality of the average face.

5.2. Face reconstruction of the broken Avalokitesvara in Beishan No. 180

As the efficient face reconstruction of the Willow Avalokitesvara in Shimenshan No.6, we enforce another reconstruction task. Shown in Fig.9, Beishan No. 180 is a cave of "Thirteen incarnations of Avalokitesvara" for there are 12 statues on left and right sides and 1 main statue in the middle. In which the outermost two statues of the cave are completely broken, and the Sta. 5 and 6 were discov-

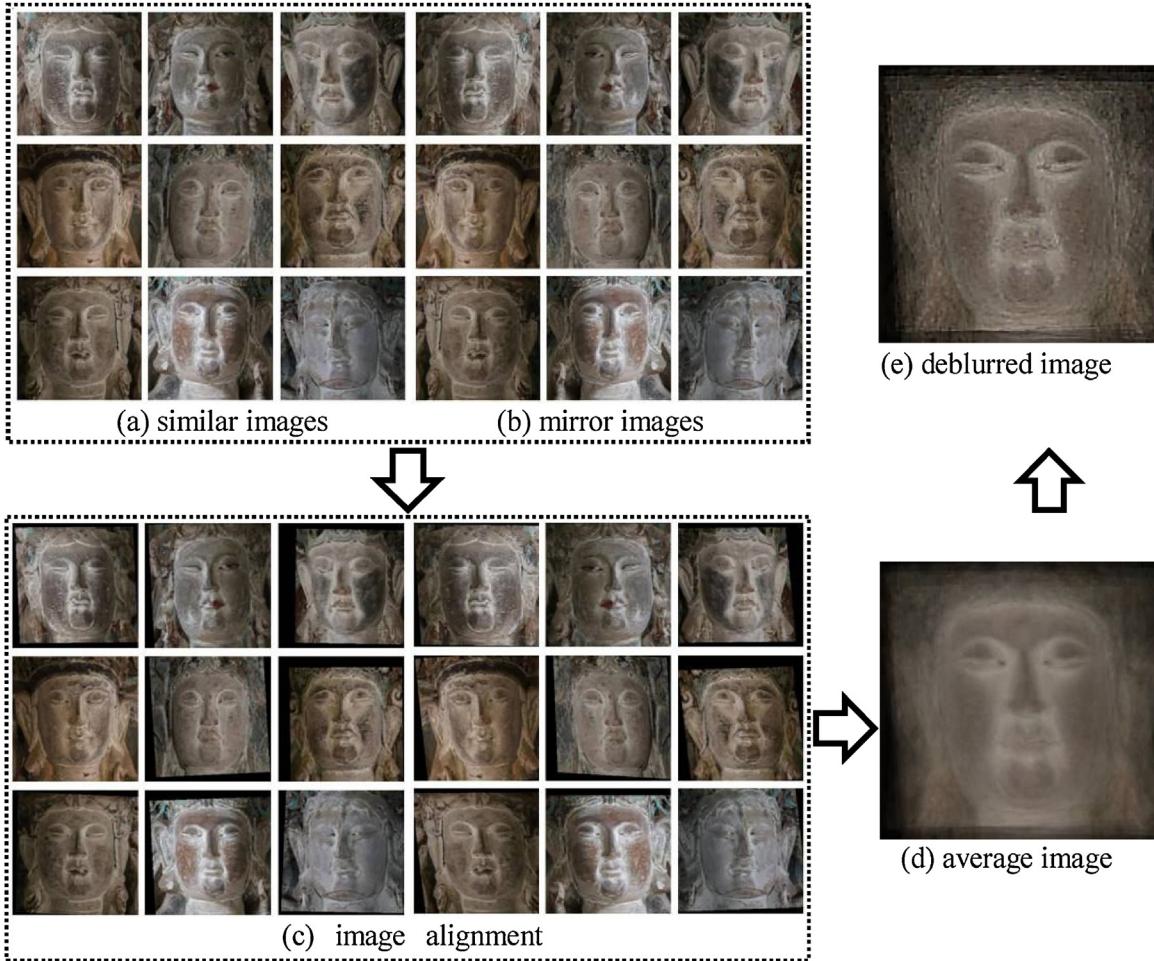


Fig. 6. Average face generation of the Willow Avalokitesvara.

Table 2

Paired similarity value between the averaged face based on different reference images.

Image	Avg(5)	Avg(6)	Avg(7)	Avg(8)	Avg(9)	Avgm(10)	Avgm(12)	Avgm(14)	Avgm(16)	Avgm(18)
Avg(5)	1	0.9887	0.9852	0.9678	0.9780	0.9783	0.9832	0.9757	0.9687	0.9645
Avg(6)	0.9887	1	0.9916	0.9802	0.9856	0.9696	0.9821	0.9749	0.9719	0.9598
Avg(7)	0.9852	0.9916	1	0.9773	0.9905	0.9822	0.9893	0.9848	0.9806	0.9763
Avg(8)	0.9678	0.9802	0.9773	1	0.9863	0.9593	0.9608	0.9704	0.9768	0.9521
Avg(9)	0.9780	0.9856	0.9905	0.9863	1	0.9786	0.9818	0.9832	0.9840	0.9804
Avgm(10)	0.9783	0.9696	0.9822	0.9593	0.9786	1	0.9906	0.9905	0.9788	0.9881
Avgm(12)	0.9832	0.9821	0.9892	0.9608	0.9818	0.9906	1	0.9900	0.9791	0.9852
Avgm(14)	0.9757	0.9749	0.9848	0.9704	0.9832	0.9905	0.9900	1	0.9934	0.9884
Avgm(16)	0.9687	0.9719	0.9806	0.9768	0.9840	0.9788	0.9791	0.9934	1	0.9841
Avgm(18)	0.9645	0.9598	0.9763	0.9521	0.9804	0.9881	0.9852	0.9884	0.9841	1

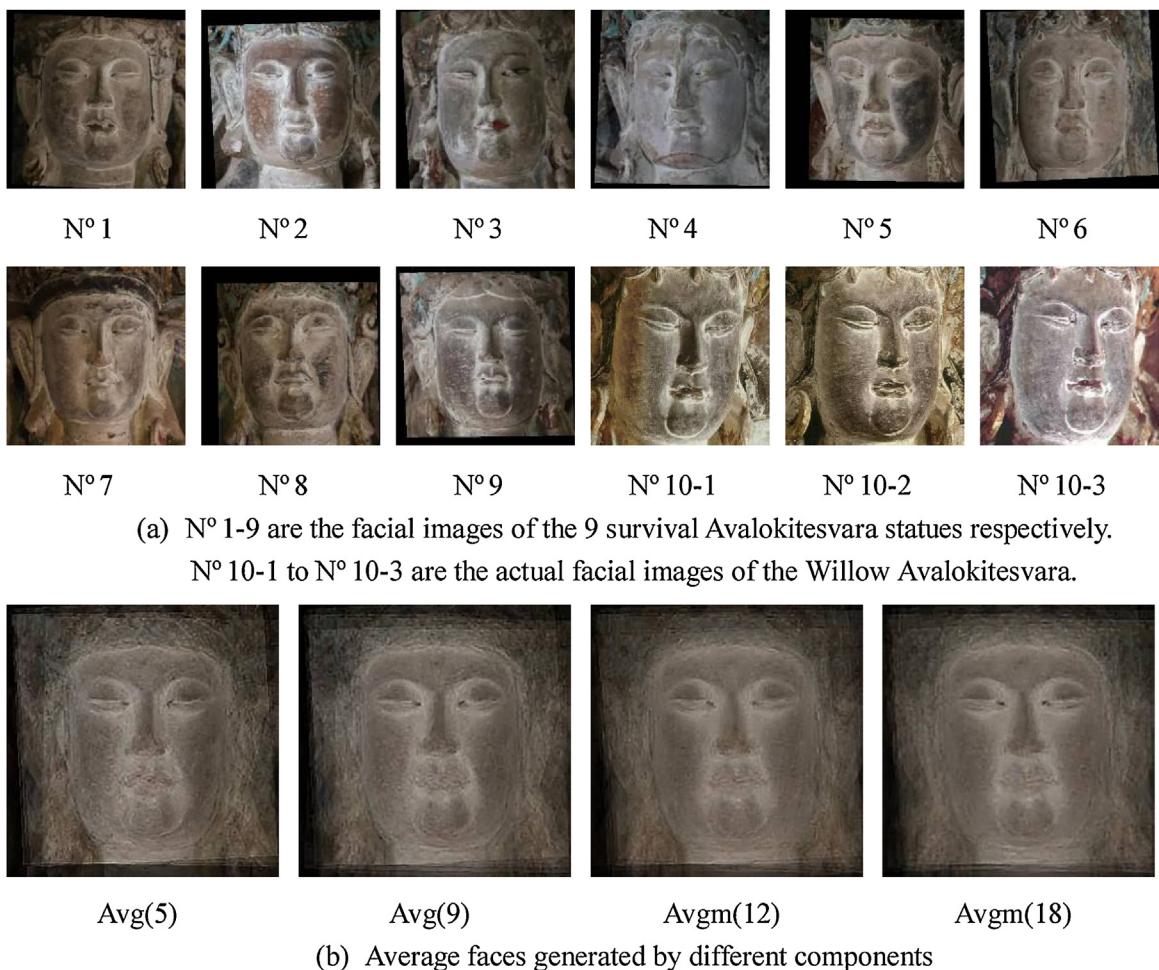
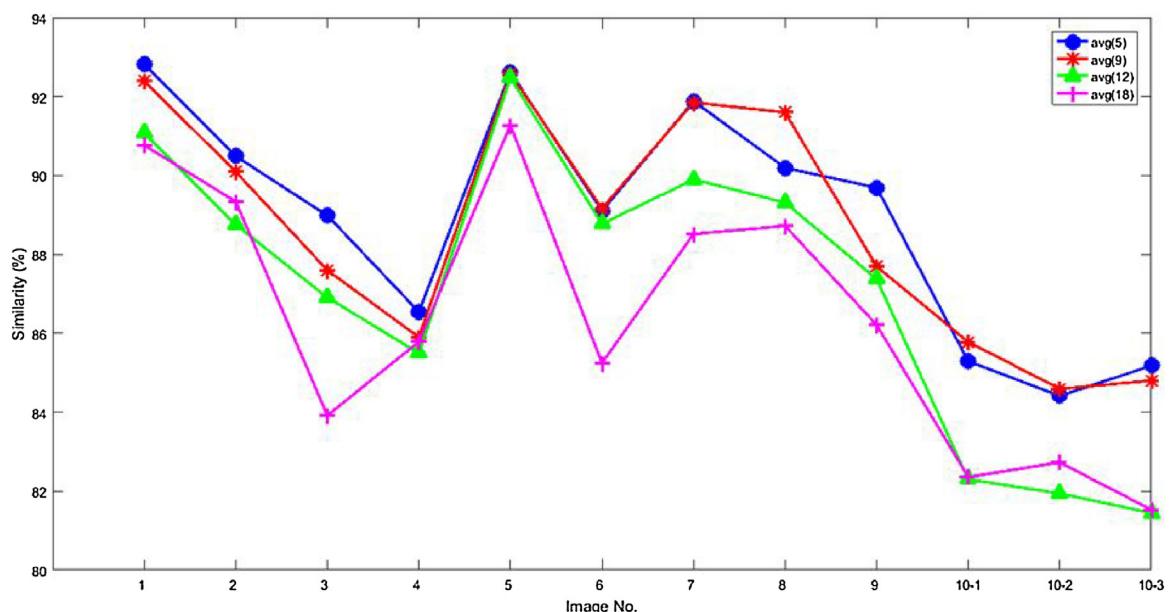
The highest (except of 1) and lowest similarity value in each column are highlighted in red bold font and blue italic font respectively.

ered to be constructed to substitute the outermost two damaged statues to keep the integrity of the cave [24].

The survival Avalokitesvara images of Sta. 1–4 and Sta. 7–10 in the cave are averaged to be two facial images as the outermost two statues' face respectively. In Fig. 9(a) and (a'), 8 front view images of the eight statues and their mirror images are directly selected as the 16 reference images. For each image, we make a face shot and warp for alignment in Fig. 9(b) and (b'). The pixels of every 8 images are averaged to create one facial image respectively. The two average faces can be viewed as the plausible faces of the outermost two broken statues after deblurring, illustrated in Fig. 9 (d) and (d'). The virtual proposals for the outermost two statues are considered reliable and acceptable by Xiaomei Wei, an assistant researcher of Academy of Dazu Rock Carvings.

5.3. Comparison

The average image is then applied to exemplar-based inpainting framework for Dazu rock carvings on the basis of a sample database. The searching region for matching patch is extended from the remainder of the original image to an average image. To validate the efficiency of our average-face-based inpainting method, we implement some experiments compared with the classical exemplar-based inpainting algorithm proposed by Criminisi et al. [5], an excellent image completion using planar structure guidance proposed by Huang et al. [39], and scene completion proposed by Hays and Efros [25]. The former two enforce inpainting within a single image (the damaged image itself), and the latter is based on multiple images.

**Fig. 7.** Comparison between the average image and the actual Avalokitesvara image.**Fig. 8.** Similarity measurement between the average image and the actual image.

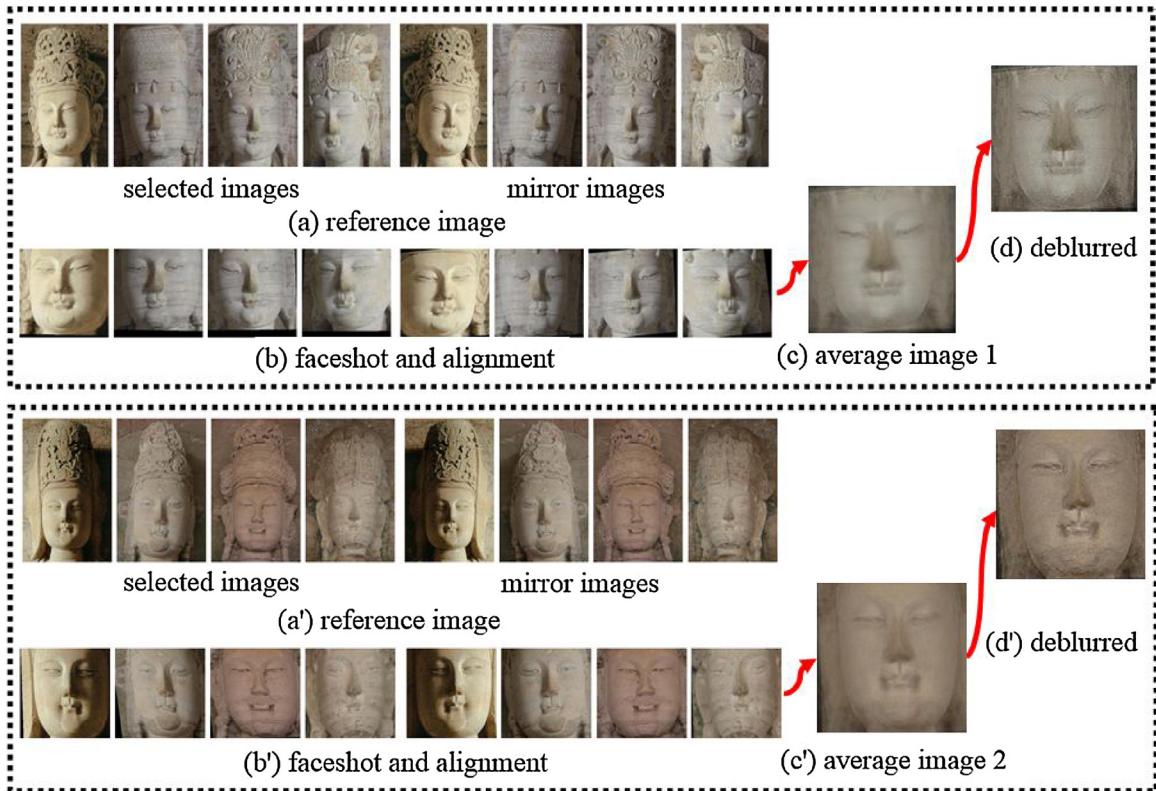


Fig. 9. Reconstruction of the faces of the broken statue in Beishan No. 180.

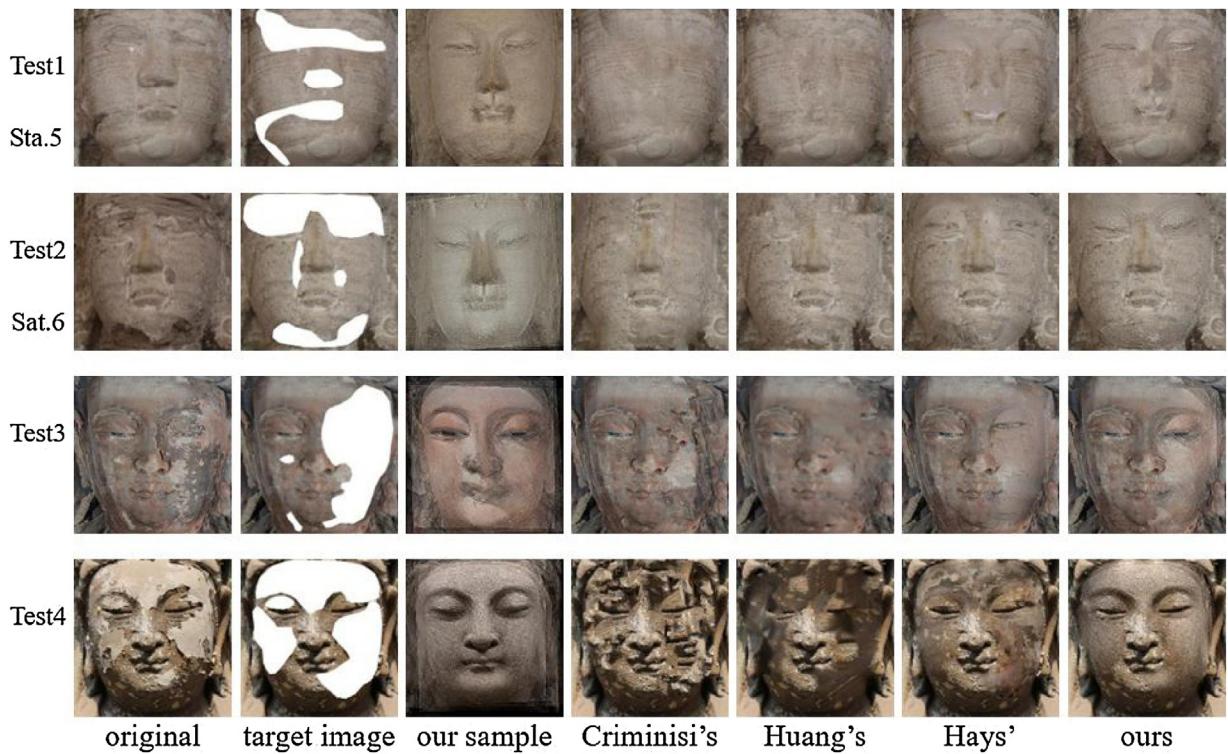


Fig. 10. Comparison of the inpainting results.

Test 1 and 2 are the facial image inpainting results of Sta. 5 and Sta. 6 in Beishan No. 180. The two average faces generated in Section 5.2 are used to be the source image respectively. According to the comparative results in Fig. 10, our average-face-based inpainting method are absolutely more efficient than Criminisi and Huang algorithms for both Sta. 5 and 6. In Test 1, both of them fail to reconstruct the facial important elements, such as eyes, nose and mouth. While our method successfully restores them. The same results are revealed in Test 2.

Compared to Hays method, ours can get competitive, sometimes better inpainting results. Illustrated in Fig. 10, there is inconsistent completion of the wing of nose for Sta. 5 in Hays' result, although the bridge of the nose reconstructs well. Whereas ours does much better for our method successfully restores both eyes and nose. For Sta. 6, ours and Hays' are competitive in vision perception.

Furthermore, we deal with another two inpainting tasks for large-scale and irregularly damaged images. Test 3 is the inpainting results for a statue in Baodingshan No. 29, and Test 4 is for a statue on Baodingshan No. 18.

It demonstrates that our method produces the most visually pleasing inpainting results. For example, Criminisi and Huang algorithms fail to reconstruct symmetry elements for Test 3 as the loss of the eye and eyebrow. And there are much disgusting artefacts and blurs in the results for Test 4, suffering from synthesis errors propagation and wrong repetitive patterns. The result of Hays method is comparable to ours but there is asymmetry completion for Test 3 due to the selected similar image is not aligned with the target image in their method. Whereas our method performs better in symmetry element inpainting and leads little block artefacts (Test 3). Our result for Test 4 looks the most pleasing in vision, with almost no disturbing artefacts. It attributes to the blending process after exemplar-based inpainting in our method.

The aforementioned experiments illustrate the efficiency of our method. It performs well in reconstructing symmetry elements as well as restoring large-scale and irregular damage. It is especially superior for the inpainting case that image priors are insufficient in the original image. Compared to the related methods, the larger the missing region is, the more efficient it is.

6. Conclusion and discussion

Averageface-based Buddhist facial image inpainting method for Dazu Rock Carvings based on a sample database is proposed in this paper. Several inpainting experiments show its efficiency and advantages when there is little prior information in the original image.

Firstly, the average image can be directly viewed as the reconstruction proposal for the extremely challenging inpainting task where there is no relative priors in the original image, such as the successful reconstruction of Willow Avalokitesvara in Shimenshan No. 6 and the two broken Avalokitesvaras in Beishan No. 180. It is valuable for art image inpainting in practice. Moreover, the average image can be also as a searching region applied to exemplar-based inpainting framework. It can get more plausible result compared with the related algorithms when the damage region is large-scale and irregular.

However, manual face landmarks is still a limitation in this paper. As the facial skin damage is usual for Dazu rock carvings' faces, the existing excellent face detection technologies cannot be directly utilized in our work. Therefore, we would like to research for an efficient face detection method for rock carvings in the future. In addition, our method is only useful for the cave in which the statues are with similar modeling style. For the cave comprised of statues carved in diverse periods, dating should be firstly judged according to archeology technology and art history knowledge, or

modeling style recognition based on deep convolutional network [24].

Acknowledgements

This work was project supported by doctoral research foundation of Sichuan Fine Arts Institute, Chongqing, China [grant number 18BS002], and graduate research and innovation foundation of Chongqing, China [grant number CYS17023].

References

- [1] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, Image inpainting, in: Proc. 27th Annu. Conf. Comput. Graph. Interact. Techn., 2000, pp. 417–424.
- [2] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, J. Verdera, Filling in by joint interpolation of vector fields and gray levels, *IEEE Trans. Image Process.* 10 (8) (2001) 1200–1211.
- [3] T. Chan, J. Shen, Local inpainting models and TV inpainting, *SIAM J. Appl. Math.* 62 (3) (2001) 1019–1043.
- [4] T. Chan, S. Kang, J. Shen, Euler's elastica and curvature-based inpainting, *SIAM J. Appl. Math.* 63 (2) (2003) 564–592.
- [5] A. Criminisi, P. Pérez, K. Toyama, Region filling and object removal by exemplar-based image inpainting, *IEEE Trans. Image Process.* 13 (9) (2004) 1200–1212.
- [6] C. Barnes, E. Shechtman, A. Finkelstein, D.B. Goldman, PatchMatch: a randomized correspondence algorithm for structural image editing, *ACM Trans. Graph.* 28 (3) (2009) 1–11.
- [7] O. Le Meur, J. Gauthier, C. Guillemot, Exemplar-based inpainting based on local geometry, in: Proc. 18th IEEE ICIP, 2011, pp. 3401–3404.
- [8] K. He, J. Sun, Statistics of patch offsets for image completion, in: European Conference on Computer Vision, 2012, pp. 16–29.
- [9] A. Bugeau, M. Bertalmio, V. Caselles, G. Sapiro, A comprehensive framework for image inpainting, *IEEE Trans. Image Process.* 19 (10) (2010) 2634–2645.
- [10] M. Elad, J.L. Starck, P. Querre, D.L. Donoho, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA), *Appl. Comput. Harmonic Anal.* 19 (3) (2005) 340–358.
- [11] J. Mairal, M. Elad, G. Sapiro, Sparse representation for color image restoration, *IEEE Trans. Image Process.* 17 (1) (2008) 53–69.
- [12] Z. Xu, J. Sun, Image inpainting by patch propagation using patch sparsity, *IEEE Trans. Image Process.* 19 (5) (2010) 1153–1165.
- [13] A.D. Rosa, A. Bonacchi, V. Cappellini, M. Barni, Image segmentation and region filling for virtual restoration of artworks, *Int. Conf. Image Process., Proc. IEEE* 1 (2001) 562–565.
- [14] R. Berns, S. Byrns, F. Casadio, I. Fiedler, C. Gallagher, F. Imai, et al., Rejuvenating the color palette of Georges Seurat's A Sunday on La Grande Jatte-1884: a simulation, *Color Res. Appl.* 31 (4) (2006) 278–293.
- [15] T. Ružić, B. Cornelis, L. Platiša, A. Pižurica, A. Dooms, W. Philips, et al., Virtual restoration of the Ghent altarpiece using crack detection and inpainting, *Lect. Notes Comput. Sci.* 6915 LNCS (2011) 417–428.
- [16] M. Fornasier, D. Toniolo, Fast, robust and efficient 2D pattern recognition for re-assembling fragmented images, *Pattern Recognit.* 38 (11) (2005) 2074–2087.
- [17] M. Fornasier, R. Ramlau, G. Teschke, The application of joint sparsity and total variation minimization algorithms to a real-life art restoration problem, *Adv. Comput. Math.* 31 (1–3) (2009) 157–184.
- [18] A. Maronidis, C. Voutounos, A. Lanitis, Designing and evaluating an expert system for restoring damaged byzantine icons, *Multimedia Tools Appl.* 74 (21) (2015) 9747–9770.
- [19] F. Stanco, G. Ramponi, A. De Polo, Towards the automated restoration of old photographic prints: a survey, *Eurocon 2003. Comput. As A Tool. IEEE Region 2* (2003) 370–374.
- [20] T. Saito, T. Komatsu, T. Ohuchi, T. Seto, Image processing for restoration of heavily-corrupted old film sequences, *Int. Conf. Patt. Recognit* 3 (2000) 13–16.
- [21] S. Tilie, L. Laborelli, I. Bloch, Blotch detection for digital archives restoration based on the fusion of spatial and temporal detectors, in: International Conference on Information Fusion, 2006, pp. 1–8.
- [22] S. Li, Chinese Bodhisattva Images, Arts and Religious Civilization of Chang'an, Zhonghua Book Company, Beijing, 2002, pp. 143–220.
- [23] Y. Lei, Overview of Sichuan Buddhist Cave Temples and Rock Carvings, Buddha, Images and Cultural Heritage, Chongqing University Press, Chongqing, 2016, pp. 95–124.
- [24] H. Wang, Z. He, Y. Huang, D. Chen, Z. Zou, Bodhisattva head images modeling style recognition of Dazu Rock Carvings based on deep convolutional network, *J. Cult. Herit.* 27 (2017) 60–71.
- [25] J. Hays, A.A. Efros, Scene completion using millions of photographs, *Commun. ACM* 51 (10) (2008) 87–94.
- [26] H. Amirshahi, S. Kondo, K. Ito, T. Aoki, An image completion algorithm using occlusion-free images from Internet photo sharing sites, *IEICE Trans. Fundamentals Electron. Commun. Comput. Sci* E91-A (10) (2008) 2918–2927.
- [27] O. Whyte, J. Sivic, A. Zisserman, Get Out of my Picture! Internet-based Inpainting, in: British Machine Vision Conference. Proceedings, 2009, pp. 116.1–116.11.
- [28] G. Wang, X. Chen, S. Hu, Geometry-aware image completion via multiple examples, *Pac. Graph.* (2014).

- [29] J.H. Langlois, L.A. Roggman, Attractive faces are only average, *Psychol Science*. 1 (2) (1990) 115–121.
- [30] L. Xu, J. Jia, Two-phase kernel estimation for robust motion deblurring, in: Computer Vision - ECCV 2010, Proceedings, 2010, pp. 157–170.
- [31] G. Rhodes, F. Proffitt, J. Grady, A. Sumich, Facial symmetry and the perception of beauty, *Psychon. Bull. Rev.* 5 (4) (1998) 659–669.
- [32] G. Rhodes, S. Yoshikawa, A. Clark, K. Lee, R. McKay, S. Akamatsu, Attractiveness of facial averageness and symmetry in nonwestern cultures: in search of biologically based standards of beauty, *Perception* 30 (2001) 611–625.
- [33] D.I. Perrett, D.M. Burt, I.S. Penton-Voak, K.J. Lee, D.A. Rowland, R. Edwards, Symmetry and human facial attractiveness, *Evol. Hum. Behav.* 20 (1999) 295–307.
- [34] T. Valentine, S. Darling, M. Donnelly, Why are average faces attractive? The effect of view and averageness on the attractiveness of female faces, *Psychon. Bull. Rev.* 11 (3) (2004) 482–487.
- [35] B.C. Jones, L.M. Debruine, A.C. Little, The role of symmetry in attraction to average faces, *Percept. Psychophys.* 69 (8) (2007) 1273–1277.
- [36] B.C. Jones, A.C. Little, D.M. Burt, D.I. Perrett, When facial attractiveness is only skin deep, *Perception* 33 (2004) 569–576.
- [37] A.C. Little, P.J. Hancock, The role of masculinity and distinctiveness in judgments of human male facial attractiveness, *Br. J. Psychol.* 93 (4) (2002) 451–464.
- [38] P. Pérez, M. Gangnet, A. Blake, Poisson image editing, *ACM Siggraph*, 22 (3) (2003) 313–318.
- [39] J.B. Huang, S.B. Kang, N. Ahuja, J. Kopf, Image completion using planar structure guidance, *ACM. Trans. Graph.* 33 (4) (2014) 1–10.