

REVIEW

Open Access

Exemplar-based image inpainting using angle-aware patch matching



Na Zhang, Hua Ji*, Li Liu and Guanhua Wang

Abstract

Image inpainting has been presented to complete missing content according to the content of the known region. This paper proposes a novel and efficient algorithm for image inpainting based on a surface fitting as the prior knowledge and an angle-aware patch matching. Meanwhile, we introduce a Jaccard similarity coefficient to advance the matching precision between patches. And to decrease the workload, we select the sizes of target patches and source patches dynamically. Instead of just selecting one source patch, we search for multiple source patches globally by the angle-aware rotation strategy to maintain the consistency of the structures and textures. We apply the proposed method to restore multiple missing blocks and large holes as well as object removal tasks. Experimental results demonstrate that the proposed method outperforms many current state-of-the-art methods in patch matching and structure completion.

Keywords: Image inpainting,, Surface fitting,, Angle awareness,, Dynamic patch selection

1 Introduction

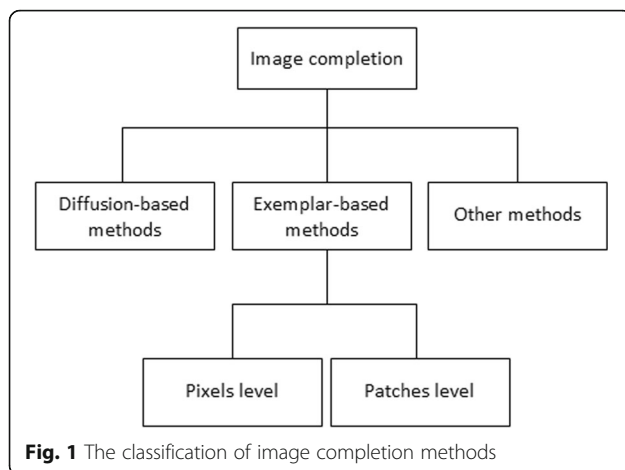
With the advancement of society and the rapid development of the Internet, image completion, also called image inpainting, has been applied to many fields proverbially, such as the protection of ancient relics, image editing, medical field, and military field. The technology is originally developed to renovate damaged photos and films or to remove unwanted texts and the occlusion from images in a plausible way, etc. All image completion algorithms are based on a hypothesis that the completed region and the missing region have the same statistical property and geometric structure. In addition, the inpainted image should satisfy the human visual consistency requirement as much as possible. That is, the colors, textures, and geometric structures of the inpainted regions should be similar to those of the ambient regions. Based on this assumption, many researchers have put forward their researches and obtained significant success. In the next paragraphs, we introduce several main completion methods and the classification diagram is shown in Fig. 1.

Currently, image inpainting methods have been mainly divided into two categories. The first category is

diffusion-based methods [1–4], where parametric models are established by partial differential equations in order to propagate the local structures from known regions to unknown regions. Bertalmio et al. [1] proposed the Bertalmio–Sa-piro–Caselles–Ballester-based inpainting method, in which the information around the defective area was propagated from the outside to the inside along the direction of the isophotes in incomplete regions, thereby obtaining the restored image from the damaged one. Motivated by the idea, Chan and Shen [2] proposed the total variation (TV) model which can restore and maintain edge information, while performing denoising by using anisotropic diffusion. But, the method merely depended on its gradient values rather than geometric information of the isophotes, which led to the applicability of small missing regions. Furthermore, Chan and Shen proposed curvature-driven diffusions (CCD) model [3] in which the diffusion process took structure information of contour (curvature term) into account. When there was a large curvature anywhere, the isophotes became strong and then gradually weakened as the isophotes extended. It suppressed the large curvature and protected the small in the restoring process so that the “connectivity criterion” was satisfied. Therefore, compared with TV model, the CCD model can repair not only images with large damaged areas but

* Correspondence: jihua@sdnu.edu.cn

School of Information Science and Engineering, Shandong Normal University, Jinan 250014, Shandong Province, China



also fine edges, especially for gray-scale images. In recent years, other diffusion-based methods have been studied for restoring images [5–8]. Biradar and Kohir [6] made use of median filter to preserve important properties of edges through diffusing median information of pixels from the outside to the inside in inpainted region. Prasath et al. [8] combined total variation with regularization to solve ill-posed image processing problems. In short, the diffusion-based methods can only restore natural and small-scale images with lower structures of texture and geometric. When containing large holes (e.g., the missing regions) or complex textures, the restored part will be over-smoothed and produce unpleasant artifacts which led to the inconsistency of structure and texture.

Therefore, the second category is exemplar-based methods which have been presented to complete an image with a large missing region [9–13]. Exemplar-based inpainting methods filled in the missing region at a pixel level or a patch level. Due to massive runtime and inconsistent texture synthesis at a pixel level, Efors and Leung [9] proposed a patch-based inpainting method to fill in the unknown region by using texture synthesis. Afterward, Criminisi et al. [10] presented a classical inpainting method to remove a large object according to the computation of priority and similarity of patches, while preserving important information of texture and structure. After that, many other exemplar-based methods have been proposed. For example, Sun et al. [11] proposed a global structure propagation method which connected the geometry structure of the whole image manually and then propagated the known texture to an unknown region. Wong et al. [12] proposed a non-local mean by utilizing multiple samples in an image to obtain more similar source patches. Based on the method achieved in [11], Li and Zhao [14] proposed an automatic structure completion method which avoided manual intervention and enhanced the efficiency of the

algorithm. In addition, the dropping effect of a confidence term was also an obvious drawback, so Wang et al. [15] introduced a regularized factor to limit the adverse effect and improved matching accuracy through a two-round search. Hereafter, a multi-scaled space method using a multi-resolution hierarchy was proposed by Kim et al. [16] and Liu et al. [17] that it can reduce workload and restore more textural information and structural features. Wang et al. [18] improved the priority estimation and structure consistency considered in patch matching. These methods were proud of accomplishing the textural and structural coherence inside the cavity.

Unlike the aforementioned exemplar-based methods, other methods were also applied to image inpainting and got satisfactory inpainted results. For example, Komodakis et al. [19] proposed the discrete global optimization tactics optimized by priority-BP based on a Markov random field. But this method was time-consuming. Subsequently, Ruzic et al. [20] presented context-aware patch-based image inpainting which divided the image into patches with variable sizes. MRF encoded consistency of neighboring patches. To fill large holes surrounded by different types of structure and texture, Alotaibi and Labrosse [21] compared with the previous methods in terms of speed and performance and obtained significant improvement. Ge et al. [22] made use of an optimal seam method to synthesize texture seamlessly and proposed patch subspace learning to limit patch selection. The problem was solved by EM-like algorithm, but it was sensitive to the initial values. Zhao et al. [23] proposed a coherent direction-aware patch alignment scheme based on GPU to enhance the similarity between matching patches, so the runtime can be reduced compared with similar methods. Huang et al. [24] proposed an automatically guiding patch-based image completion which considered patch transformation and used the mid-level constraints to guide the filling process.

In this paper, our novel scheme is performed in a dynamic manner. We initialize the unknown regions of a damaged image with the help of the MLS method. According to a dynamic patch selection process, small target patches are applied in the high-frequency region to maximize the restoration of the structural information, while large patches can reduce the computational workload in the low-frequency region. By developing the angle-aware patch matching with a Jaccard similarity coefficient, the process of patch matching can be performed better to advance matching accuracy. And the proposed approach can select multiple matching patches automatically from the available region by using an angle-aware rotation strategy to increase the probability of obtaining the optimal matching patch. It is worth

mentioning that we also improve the priority function based on [10] for the selection of the target patch. In summary, the main contributions of this paper are as follows:

1. The proposed algorithm is first initialized by a surface fitting technique using the moving least squares method. Then, the estimated values can be viewed as the prior knowledge for the following inpainting.
2. We capture the structural details efficiently by using different sizes of the target patches and reduce computational time.
3. A novel patch-matching algorithm is proposed in which we introduce a Jaccard similarity coefficient that can improve the matching precision between patches. Meanwhile, we also add the gradient information to the matching process. The most important point is that we rotate the direction of matching patches to find an optimal result.
4. Finally, we compare it with state-of-the-art image inpainting approaches on some real images. The final results illustrate the superiority of the proposed method in terms of accuracy and efficiency.

The remainder of this paper is organized as follows. Section 2 briefly narrates the classical exemplar-based method. Section 3 introduces the proposed approach in detail. Section 4 compares the proposed method with other state-of-the-art methods in some real images. The proposed method is compared with deep learning for image inpainting in Section 5. Finally, we draw a conclusion for this paper in Section 6.

2 The general exemplar-based inpainting

In exemplar-based image inpainting methods, the most famous method was proposed by Criminisi et al. [10], which can restore structural and textural information of the large damaged region simultaneously. Given an image I decomposed into two parts as shown in Fig. 2, the source region (pixels are known) and the target region (pixels are unknown), are represented by Φ and Ω respectively, where $\Phi = I - \Omega$. $\delta\Omega$ denotes the boundary line connecting Φ and Ω together. Since the point filled earlier can affect the points filled afterward [18], we need to determine the priority of every point $p \in \delta\Omega$ by a priority function. The process is described as follows briefly. Firstly, a target patch Ψ_p centered at a point p with the highest priority is selected preferentially and it contains both the known pixels and the unknown pixels. Secondly, a source patch Ψ_q centered at a point $q \in \Phi$ is selected to match the above target patch. Finally, the

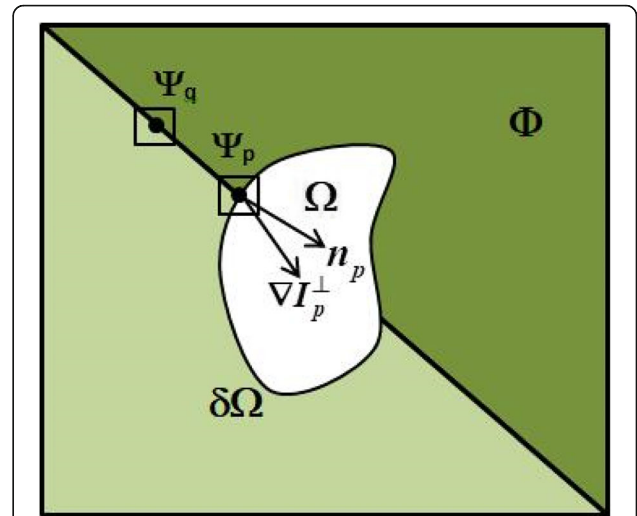


Fig. 2 Typical schematic diagram by exemplar-based inpainting. An original image I including the missing region Ω and the source region Φ . $\delta\Omega$ is the boundary line of Ω and $\delta\Omega \in \Phi$. Ψ_p is a target patch centered at the pixel $p \in \Omega$ and is a source patch Ψ_q centered at a pixel $q \in \Phi$. n_p is a unit vector orthogonal to $\delta\Omega$. ∇I_p^\perp is the isophote vector at point p

defective region is filled by copying the corresponding pixels in Ψ_q to the unknown pixels within Ψ_p .

When filling the missing cavity, we need first select the target patch to be filled with the highest priority along the boundary $\delta\Omega$. That is, the higher the priority of a patch is, the more important the contained information is. Thus, the priority function, to compute which patch will be earliest filled, is defined as

$$P(p) = C(p)D(p), p \in \delta\Omega \tag{1}$$

where $C(p)$ is the confidence term to measure the ratio of the reliable pixels in Ψ_p ; $D(p)$ is the data term to denote the strength of the isophotes along the boundary $\delta\Omega$. If a patch is along the direction of the isophotes, it will have a higher priority to be filled. The confidence term $C(p)$ is computed as

$$C(p) = \frac{\sum_{k \in \Psi_p \cap \Phi} C(k)}{|\Psi_p|} \tag{2}$$

$$C(p) = \begin{cases} 0, & \text{if } \forall p \in \Omega, \\ 1, & \text{if } \forall p \in \Phi. \end{cases} \tag{3}$$

where k is one of the common pixels in both Ψ_p and Φ . $C(k)$ is the confidence value of the pixel k . $|\Psi_p|$ represents the number of all points in the Ψ_p . And the confidences of all points in source region Φ are initialed as 1, while others are zero. Next, the data term is defined as

$$D(p) = \frac{|\nabla I_p^\perp \cdot n_p|}{\alpha} \tag{4}$$

where \perp and ∇I_p are an orthogonal operation and the gradient centered at pixel p respectively. Thus, ∇I_p^\perp denotes the isophotes vector orthogonal to the gradient ∇I_p , and n_p is a unit vector orthogonal to the boundary $\delta\Omega$. α is a normalization factor ($\alpha = 255$ if I is a gray image). Finally, through calculating the priority function, we can find the point \hat{p} with the highest priority, and construct the target patch centered at it.

After determining the target patch to be filled, we search for the best matching patch Ψ_q from the source region Φ to complete the unknown pixels of the target patch $\Psi_{\hat{p}}$. How to select a source patch or how to measure the similarity correctly between the source patch and target patch is critical for the quality of inpainted images. Criminisi's method measures the textural similarity distance between Ψ_q and $\Psi_{\hat{p}}$ by the sum of squared differences (SSD), which can be expressed as

$$D(\Psi_q, \Psi_{\hat{p}}) = \sum_{k \in \Psi_{\hat{p}}} [\Psi_q(k) - \Psi_{\hat{p}}(k)]^2 \tag{5}$$

Next, we fill the unknown pixels in the target patch $\Psi_{\hat{p}}$ according to the corresponding pixels within the most similar source patch $\Psi_{\hat{q}}$. The confidence values need also be initialized for the newly filled pixels, defined as

$$C(k) = C(\hat{p}), \forall k \in \Psi_{\hat{p}} \cap \Omega \tag{6}$$

After filling $\Psi_{\hat{p}}$, the boundary $\delta\Omega$ is updated iteratively until the unknown region is filled entirely.

3 Proposed method

We propose an exemplar-based image inpainting algorithm using angle-aware patch matching which is used to recover missing regions consisting of textural and structural components. And it can make inpainting result look more natural in connection. The overall architecture of our inpainting system is shown in Fig. 3. The first step is to initialize all unknown pixels in the missing region by surface fitting technique. At the second step, we need to calculate the priority function to determine the filling order of every pixel point at the boundary and select the target patch to be filled according to the size of gradient value of filled points dynamically. Next, we search for multiple matching patches using angle rotation strategy from the source region, and these patches have the most similar features to the target patch. And according to the proposed similarity metric, we find that the optimal source patch can achieve satisfactory inpainted results.

3.1 Initialization by surface fitting method

The aim of the subsection is to estimate pixel values for the missing region of an image. These values are plausible but have some randomness. Based on this case, we apply the surface fitting technique in 3D subspace to initialize the pixel values within the missing region. We utilize the moving least squares (MLS) method [25] to fit a surface in 3D subspace according to surrounding pixels of the missing region.

Given an image I viewed as a 2D matrix, we project pixels of the image to a 3D subspace according to similar geometrical structure and regard the gray value of each pixel as the height of the 3D coordinate. The missing pixels of the image form the holes in 3D point clouds as shown in Fig. 4a. Figure 4a is the point cloud of the incomplete image in which the black part is the missing region in 3D subspace and Fig. 4b is the fitted surface by MLS. By this initialization, we can preserve

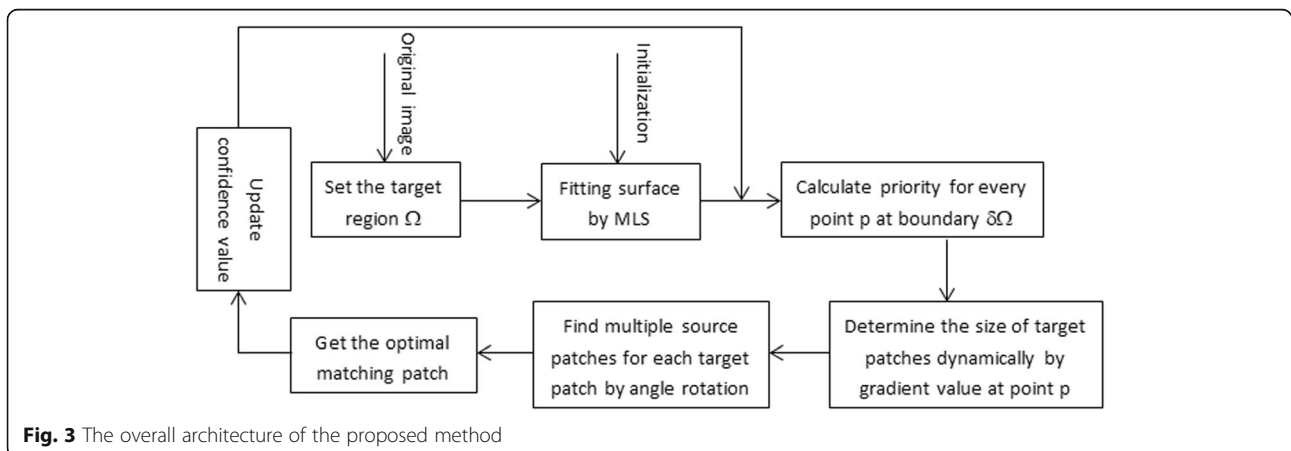


Fig. 3 The overall architecture of the proposed method

some structure features of the damaged regions. Thus, we fill the hole by fitting a surface which is generated by moving least square method and we can get a complete image I with the estimated pixels.

In Fig. 5, we fit a real color image called ‘Lena.’ Figure 5a is an original image. Figure 5b is an incomplete image with a small missing block. In Fig. 5c the hole is filled by quadric fitting. We can observe that the intensity of the fitted region is similar to the intensity of surrounding pixels. And the estimated pixels can be regarded as the prior knowledge for inpainting and provide certain structural information. More importantly, we restore the damaged region more precisely by using the proposed algorithm.

3.2 Calculation of the target patch priorities

For all points belonging to Ω , the points filled earlier can influence the points filled afterward. Thus, how to determine the filling order of these points is extremely important. The priorities of the target patches centered at these points are also critical to preserving structural information in the inpainting process. The confidence values decrease too rapidly in terms of Eq. (2) so that the priority order becomes insignificant. Recently, Wang et al. [18] propose a novel inpainting algorithm based on space varying updating strategy and structure consistent patch matching which are used to deal with the dropping problem of the confidence and improve matching quality, respectively. In this method, instead of initializing the confidences of newly filled pixels to the same value as in Eq. (6), they consider that the priority of the center pixel \hat{p} is higher than that of its surroundings. Consequently, an upper bound and a lower bound are defined to restrain the space varying confidence of points in $\Psi_{\hat{p}} \cap \Omega$. However, the values of confidence in this method only pay attention to the known pixels in

$\Psi_{\hat{p}}$ while ignoring the effect of the unknown pixels in $\Psi_{\hat{p}}$.

In our algorithm, the values of the unknown pixels have been estimated roughly as presented in Section 3.1. Thus, we need consider the contribution of all pixels in $\Psi_{\hat{p}}$. However, the estimated pixels are not enough precise in the unknown region. So we introduce weight factors for balancing the importance of pixels between $\Psi_{\hat{p}_\phi}$ and $\Psi_{\hat{p}_\Omega}$, where $\Psi_{\hat{p}_\phi}$ and $\Psi_{\hat{p}_\Omega}$ denote the known and unknown regions in $\Psi_{\hat{p}}$ respectively. A large weight should be assigned to $\Psi_{\hat{p}_\phi}$, in contrast, a small weight should be assigned to $\Psi_{\hat{p}_\Omega}$ in order to better preserve the structural information in the utmost extent. To maintain the newly filled pixels having smaller confidences than the current existing pixels, the upper bound is rewritten as

$$C_{up} = \frac{\sum_{k \in \Psi_{\hat{p}}} C(k)}{|\lambda_1 \Psi_{\hat{p}_\phi} + \lambda_2 \Psi_{\hat{p}_\Omega}|} \tag{7}$$

where $|\lambda_1 \Psi_{\hat{p}_\phi} + \lambda_2 \Psi_{\hat{p}_\Omega}|$ denotes the number of all pixels including both known pixels and estimated pixels. λ_1 and λ_2 are balancing factors which can control the status of the confidence, and they are adjusted dynamically in range (0, 1), where $\lambda_1 = 1 - \lambda_2$ and $\lambda_1 > \lambda_2$. The lower bound is set as $C(\hat{p})$.

The new confidence term can be summarized as

$$C_n(k) = \max(-\beta * dis(\hat{p}, k)^2 + C_{up}, C(\hat{p})), \quad k \in \Psi_{\hat{p}_\Omega} \tag{8}$$

where $dis(\hat{p}, k)$ is Euclidean distance between two pixels used to differentiate pixels in different locations of the patch, and β is the decreasing factor for limiting the

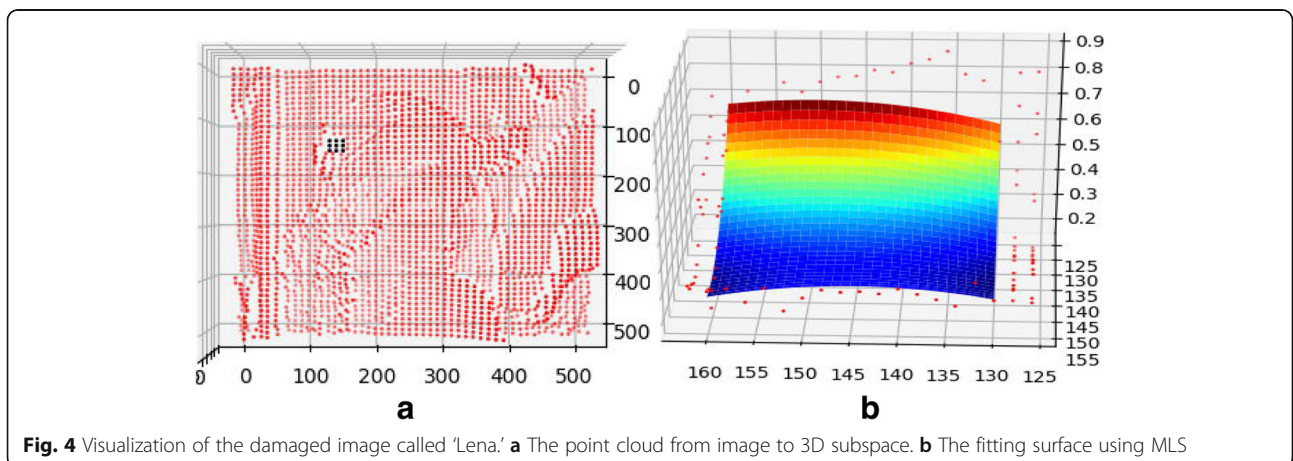


Fig. 4 Visualization of the damaged image called ‘Lena.’ **a** The point cloud from image to 3D subspace. **b** The fitting surface using MLS



Fig. 5 Fitting a missing region using MLS. **a** Original image. **b** Mask. **c** The fitted image

dropping rate of the confidence term, set as 0.02 empirically in our experiments.

The data term $D(p)$ is a benefit to reconstruct the local linear structure and texture. However, the priority value may be closer to zero when the data term is zero. In this case, in order to eliminate the above malpractice, we add a curvature factor to the data term based on [26]. Hence, $D(p)$ can be rewritten as

$$D(p) = D(p) + 1/S(p) \quad (9)$$

$$S(p) = \nabla \cdot \left[\frac{\nabla I_p}{|\nabla I_p|} \right] \quad (10)$$

where $S(p)$ is the curvature of the isophotes through the center pixel p , which produces a better effect with a significant change in the linear structure. Besides, we take a patch along the direction of the isophotes with a higher data-term value into consideration. And we introduce the intensity information of point p at the image I as $I(x_p, y_p)$, in which (x_p, y_p) is the coordinate of point p and the propagation of intensity is along the direction of the isophotes.

The priority function defining the optimal filling order can be rewritten as

$$P(p) = C(p) \left(D(p) + \frac{1}{S(p)} \right) \quad (11)$$

Finally, we find a pixel \hat{p} in the contour $\delta\Omega$ with the highest priority according to Eq. (11).

After finding a point \hat{p} , we need to construct the target patch $\Psi_{\hat{p}}$ centered at it. In previous methods, the size of the target patch is fixed, so the runtime and matching inaccuracy are increased. Based on above reasons, we introduce a novel idea that the size of the target patch should be selected dynamically according to frequency information of the image content. We notice that the high-frequency region contains more edge details and structural information, while the low-frequency region represents the smooth part in images. Therefore, we

firstly divide an image into two components: the low-frequency component R_L and the high-frequency component R_H . To speed up convergence and enhance global consistency, a small patch is used in R_H to enhance the restoration of the edge and structure details, while a large patch is employed in R_L to reduce runtime. Here, we apply a threshold operator γ set as 0.3 empirically to determine the size of each target patch. We compare threshold value γ with the gradient value of point \hat{p} . If the gradient value is larger than γ , we will construct a smaller target patch, otherwise construct a larger one. The equation of the gradient is as follows

$$\begin{aligned} G_{\hat{p}}(x, y) &= \sqrt{g_x^2 + g_y^2} \\ g_x(i, j) &= g(i+1, j) - g(i, j) \\ g_y(i, j) &= g(i, j+1) - g(i, j) \end{aligned} \quad (12)$$

where g_x and g_y denote gradients of horizontal and vertical directions of point (i, j) respectively.

3.3 Finding the optimal source patch by angle-aware patch matching scheme

At the moment, the optimal source patch should be found from the whole source region after defining the size of the target patch to be filled first. The similarity metric is very important to find the most similar patch from the source region. Previously, many traditional similarity metric methods, including Euclidean distance, mean squared error (MSE) as well as the sum of squared difference (SSD), etc., fail to adequately consider the matching coherence and ignore the differences of direction and structure variations within two patches. Due to the difference of the angle of the patterns, the same pattern can give rise to the unreasonable matching results between patches. Hence, we propose a similarity metric based on the angle-aware to avoid the inexact matching caused by different angles of patterns in two patches. In our algorithm, the purposed of rotation is to find more similar contents which ensure the consistency of matching results. With regard to this, we need to search for

multiple source patches and select the best one after rotating as shown in Fig. 6. Meanwhile, we introduce a Jaccard similarity coefficient to enhance the similarity between patches, and when we calculate the similarity between the target patch and each rotated source patch, its value is fixed. The similarity metric is defined as

$$D(\Psi_{\hat{p}}, \Psi_q) = \tau \|G_{\sigma} \otimes (\Psi_{\hat{p}} - \mathbf{R}(\Psi_q, \Delta\theta))\|^2 + \eta D_{SSD}(\nabla \Psi_{\hat{p}}, \nabla \Psi_q), \tag{13}$$

where τ is a Jaccard similarity coefficient defined as $\frac{|\Psi_{\hat{p}} \cap \Psi_q|}{|\Psi_{\hat{p}} \cup \Psi_q|}$, where $|\Psi_{\hat{p}} \cap \Psi_q|$ and $|\Psi_{\hat{p}} \cup \Psi_q|$ are the number of same pixels and the number of all pixels within $\Psi_{\hat{p}}$ and Ψ_q respectively. The larger the coefficient is, the more similar two patches are. G_{σ} is a Gaussian filter in which σ is a standard deviation whose change interval is [0.4, 0.6] leading to good inpainting results, and \otimes is the convolution operator. When $\sigma > 0.6$, the difference between the target patch and a source patch will be over-smoothed by Gaussian filter, which will yield more matching errors. When $\sigma < 0.4$, the result is poor. $\mathbf{R}(\cdot, \cdot)$ is a rotation function which rotates each source patch to guarantee the consistency of the patterns in matching results. $\Delta\theta$ is a rotated angle for each source patch, its value is 20° and the rotated range is from -90° to 90° . The purpose of this choice is that if the rotated angle is too large or too small, the inaccurate matching results will be amplified. We yet apply gradient features between two patches to the proposed distance metric in addition to color features. D_{SSD} is the sum of squared differences over the gradients of two patches, and the gradient dimensions are weighted by η .

The similarity metric considers the property of textures and structures and also merges the gradient information to make the edge more outstanding. Unlike the previous methods finding a source patch from the source region, we want to find n nearest neighbors for each target patch. In this search process, we utilize a nearest neighbor field (NNF) [23] defined as a multi-value function $f(\cdot)$ which can map each target patch coordinate to multiple source patches coordinates so that

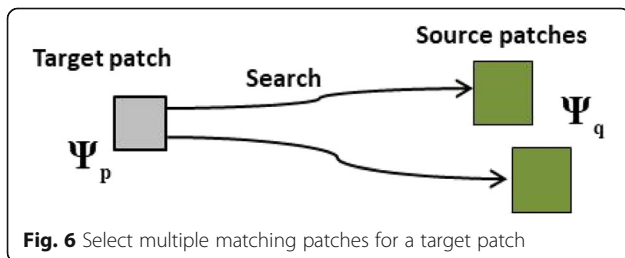


Fig. 6 Select multiple matching patches for a target patch

the matching results is more accurate. The multi-value mapping is as follows

$$f(\Psi_{\hat{p}}) = \Psi_{q_i}, i = 1, 2, \dots, n \tag{14}$$

We store these distance values between $\Psi_{\hat{p}}$ and Ψ_{q_i} to an additional array. Then, we compare these distance values by a competitive mechanism according to Eq. (13) in which we select the optimal source patch for each target patch. That is, the distance value is the smallest.

3.4 Updating the pixels of target patch

So far, we have found the best matching patch $\Psi_{\hat{q}}$ for the target patch $\Psi_{\hat{p}}$. In the last step, the previous methods directly copy the intensity values of those pixels within $\Psi_{\hat{q}}$ to corresponding pixels of the unknown part of the target patch [13, 15, 27]. In contrast, we add the intensity values of those pixels in the optimal matching patch to corresponding initialized pixels in the unknown part of the target patch, and their average values are used to fill the missing region within the target patch, which can reduce the inconsistency of structures and textures. Then, we update the confidence values of the newly filled pixels as

$$C(k) = C_n(k), \forall k \in \Psi_{\hat{p}\Omega} \tag{15}$$

The boundary $\delta\Omega$ is also updated. And the total process is repeated until the final result is obtained.

4 Results and discussion

In this section, we test the proposed approach on all kinds of natural and textural images which are selected from references as well as the Berkeley image dataset. We compare our approach with other inpainting methods including Criminisi's algorithm [10], image melding using patch-based synthesis by Darabi et al. [28], image restoration via group-based sparse representation (GSR) by Zhang et al. [29], and annihilating filter-based low-rank Hankel matrix approach (ALOHA) proposed by Jin and Ye [30]. Our experiments are simulated using Matlab 2016b in WIN10 system with Intel(R) Core(TM) i5-4590 CPU (3.30 GHz). We apply our algorithm to inpainting of missing blocks and object removal task for color images. And our approach can also be applied to a grayscale image. In our experiments, the parameters are set as follows: λ_1 and λ_2 is set as 0.7 and 0.3 respectively. Furthermore, we search for $n = 3$ the most similar source patches for every target patch. And, we set η as 0.3 in Eq. (13) to balance the contribution of the

gradient features. We set the search range as $[-\frac{\pi}{2}, \frac{\pi}{2}]$ for rotation. The approach is fairly robust in the variation of this range.

4.1 Image quality analysis methods

To evaluate the objective quality of the inpainted images we use the peak signal-to-noise ratio (PSNR) [31] and structure similarity index (SSIM) [32]. The PSNR is defined as follows

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE(I, \hat{I})} \right) \quad (16)$$

$$MSE = \frac{1}{m * n} \sum_{i=1}^n \sum_{j=1}^m \|\hat{I}(i, j) - I(i, j)\|^2 \quad (17)$$

in which the value of MAX is 255 representing the largest gray value of image color, and $m * n$ denotes the number of pixels. I indicates the original complete image, and \hat{I} is the inpainted image. In general, the larger the PSNR value is, the lesser the diversity is between the original image and the inpainted one.

Furthermore, SSIM is used for measuring the similarity between two images. The SSIM is expressed as follows

$$SSIM(I, \hat{I}) = \frac{(2\mu_I\mu_{\hat{I}} + c_1) + (2\sigma_{I\hat{I}} + c_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + c_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + c_2)} \quad (18)$$

where μ_I is the average of I , $\mu_{\hat{I}}$ is the average of \hat{I} ; σ_I^2 is the variance of I , and $\sigma_{\hat{I}}^2$ is the variance of \hat{I} ; $\sigma_{I\hat{I}}$ is the covariance of I and \hat{I} ; $c_1 = (k_1L)^2$ and $c_2 = (k_2L)^2$ are the two variables to stabilize the division with weak denominator; L is the dynamic range of the pixel values (typically this is $2^{\#bits \text{ per pixel}} - 1$); $k_1 = 0.01$ and $k_2 = 0.03$ by default [32].

In the next subsections, we will discuss the visual effectiveness of our approach and its runtime. And we also discuss the feasibility of our method in image compression.

4.2 Recovery of multiple block losses and large holes

We first apply our approach to restore the damaged images with multiple missing blocks. We compare our approach to Crinimisi's algorithm [10] and the experimental result is shown in Fig. 7, in which nine missing blocks are all set to 20×20 . We can observe clearly that her face is mottled and the stripes on the headscarf are also blurred in Fig. 7c. That is, Crinimisi's algorithm fails to recover textures and structures of missing regions. However, Fig. 7d displays the inpainted result of our approach which can restore both complex textural and

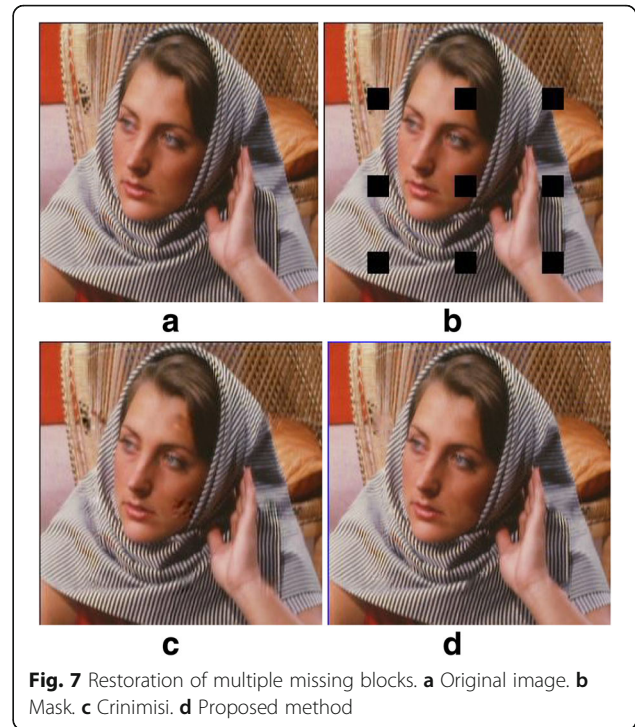


Fig. 7 Restoration of multiple missing blocks. **a** Original image. **b** Mask. **c** Crinimisi. **d** Proposed method

structural details. In contrast, our method can produce a satisfactory visual effect.

To further demonstrate the advancement of our approach, we give average image recovery accuracy according to PSNR and SSIM on a dataset of thirty images with nine missing blocks taken from the Berkeley image dataset. In Fig. 8, we compare the average PSNR and SSIM of every missing block with other methods respectively. From Fig. 8a, b, we notice that average PSNR and SSIM of the proposed algorithm are higher than those of other methods in a majority of blocks. When missing regions are smooth, Crinimisi's and GSR algorithms can also obtain satisfactory and natural effects. But if the missing regions have complex textures and structures, these regions will be filled using error contents. In addition, ALOHA's result is generally satisfactory, but it is still slightly inferior to our method. Since the proposed method restores the damaged regions using dynamic patches whose sizes are determined by gradient values, our method makes the inpainted image more reasonable in visual, and decreases the run-time. When gradient value is smaller than γ , we set the size of the target patch as 5×5 , otherwise set as 7×7 . Therefore, according to the above experimental results, we can prove that the proposed method is more effective for images with rich textures and structures.

Then our method is also suitable for filling a large hole inside images. For example, in Fig.9 we mainly give the inpainting results about filling a large hole set as 70×60 in an image of size 256×256 . Fig.9 (a) is an original

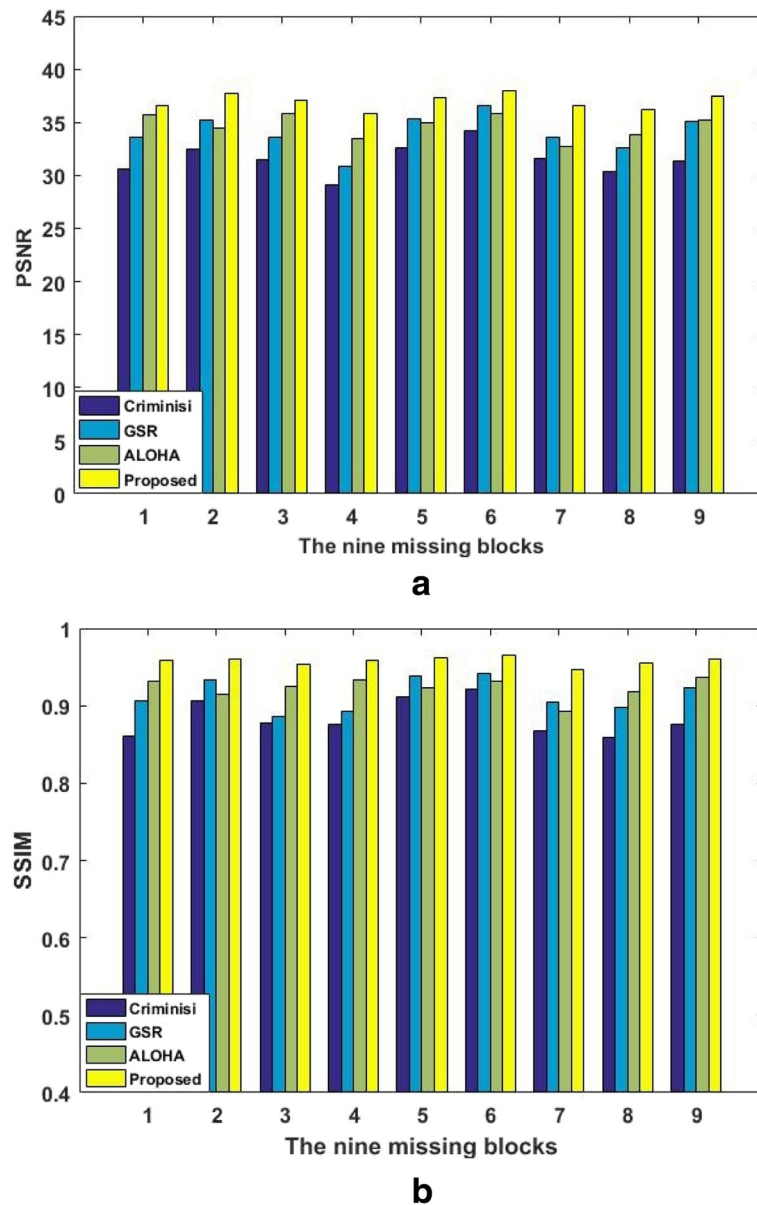


Fig. 8 Comparing results of the average PSNR and SSIM of test images with nine missing blocks. **a** The average PSNR of test images with nine missing blocks. **b** The average SSIM of test images with nine missing blocks.

image, and Fig. 9 (b) is the masked image with a target region of a rectangular block. Figs. 9 (c)-(g) are the inpainting results of the Criminisi algorithm, GSR, ALOHA, image melding, and the proposed algorithm. Clear errors can be observed in Figs. 9 (c)-(e) in which the textures of the butterfly wing are very different from the original one. From Fig. 9 (f) we observe that the inpainting result is better than the previous methods in terms of inpainting textures, but it is slightly inferior to ours as shown in Fig. 9(g). Therefore our approach can better restore not only a large hole but also textural and structural information.

4.3 Object removal

We also test our approach on the object removal task in Fig. 10. The performance of the object removal task relies on the two main aspects: image naturalness and the runtime. In this subsection, we show the results of the various algorithms on a variety of color images of size 383×256 by discussing these aspects. The removed objects in these images are shown in Fig. 10 row 2 using black block masks. From the first image in row 3, we notice that the cloud generated by Criminisi's algorithm is not smooth and not natural enough, and the connection with the rainbow also produces the mottled content. In

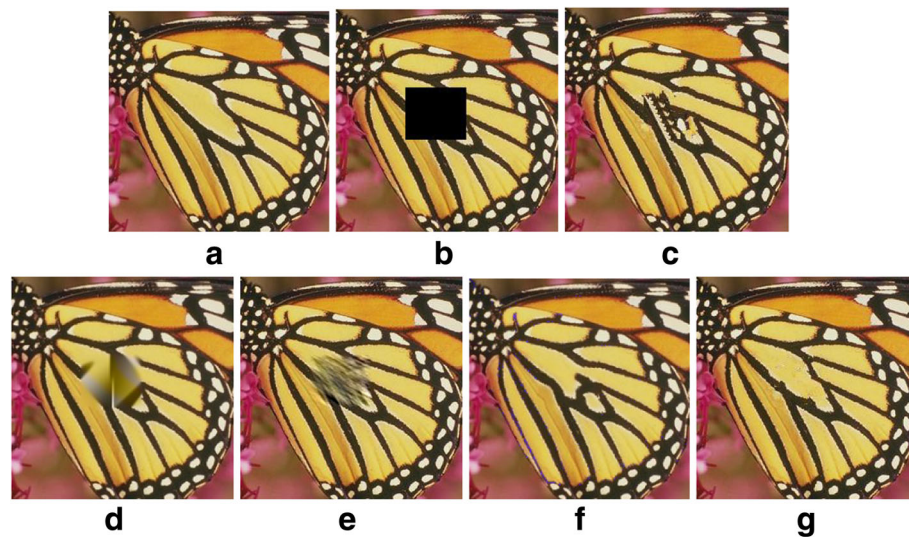


Fig. 9 The filling of a large hole. **a** Original. **b** Mask. **c** Crinimisi. **d** GSR. **e** ALOHA. **f** Melding. **g** Ours

contrast, the proposed algorithm generates relatively natural contents compared with the original image. As we have seen in row 4, GSR-based algorithm gives rise to the most unsatisfactory result, and we can clearly observe that the inpainted regions have serious inconsistency with surrounding contents. From Fig. 10 row 5 of column 2, we observe that the reflection of a person has not been reconstructed by Melding algorithm and is discontinuous. However, our algorithm and Crinimisi's algorithm can both restore the shadow better. In column 3, these methods yield the severely blurred contents in the auditorium. Overall, compared with these methods, our method generates more plausible details without sacrificing the naturalness. From these results, we discover that the restoration of texture and structure plays an important role in the whole inpainting process. By comparing with other methods, experimental results prove that our method can complete more realistic contents in visual.

In Section 3.2, we proposed that the size of the target patch is selected dynamically. The patch size is a pivotal parameter in patch-based algorithms. Larger patches capture more structures and edge details, if good matches are found. However, if such matches are not found, the result can easily converge to a blurry solution. Thus, we suggest that if the gradient of point \hat{p} with the highest priority is higher than γ , we set the size of the target patch centered at the point \hat{p} as 5×5 . Otherwise, it is set as 10×10 . If patches are too large, error contents will be yielded. However, smaller patches may consume much runtime. Therefore, we exploit dynamic patches to preserve the structures as well as strong edges in transitional areas, meanwhile, to advance the performance efficiency of our

approach. Consequently, Fig. 11 gives the comparison between a fixed patch set as 7×7 and dynamic patches, where we mark the inpainted roof with a red bold square. Obviously, the chasm is generated on the inpainted roof and structures are discontinuous as shown in Fig. 11b. In contrast, the inpainted result using dynamic patches is more acceptable and reasonable in visual in Fig. 11c. To prove the advantage of dynamic patches, we compare the mean runtime of our approach for 30 images with Crinimisi's method, GSR, image melding, and ALOHA on the Berkeley image dataset shown in Table 1. We find that our approach is faster than other inpainting methods except for Crinimisi's method. This is mainly because we use the Gaussian convolution in the computation of the similarity metric and search for multiple source patches for each target patch.

4.4 Image compression

Image compression is a technology that represents the original pixel matrix with fewer bits in a lossy or lossless manner, also known as image coding, and is the application of data compression technology on digital images. The purpose is to reduce redundant information in image data and store and transmit data in a more efficient format. In order to guarantee communication efficiency and save network bandwidth, compression techniques can be implemented on digital content to reduce redundancy, and the quality of the decompressed versions should also be preserved. Nowadays, most digital contents, particularly digital images and videos, are converted into the compressed forms for transmission [33–36]. Recently, many image inpainting methods have been applied to the image and video compression



Fig. 10 The restore results of object removal task. Row 1, Original images. Row 2, Masked images. Row 3, Criminisi's method [10]. Row 4, GSR method [29]. Row 5, Image melding method [28]. Row 6, ALOHA method [30]. Row 7, Proposed method

tasks. For example, Liu et al. [33] proposed a compression-oriented edge-based inpainting algorithm in which more redundant information will be removed during encoding for increasing the compression ratio due to adopting assistant information; then the removed regions can be restored at the decoder side to achieve good visual quality. Qin et al. [34] proposed a novel joint data-hiding and compression scheme using side match vector quantization and PDE-based image inpainting,

which is applied to the decompression process successfully. Currently, Qin and Zhou et al. [35] proposed a novel lossy compression scheme for an encrypted image in which the final reconstructed image can be generated with the help of image inpainting based on a total variation model. Besides, an inpainting algorithm based on de-interlacing method was applied to a video processing task by Coloma et al. [36] which view the lines to interpolate as gaps to be inpainted. Inspired by the above

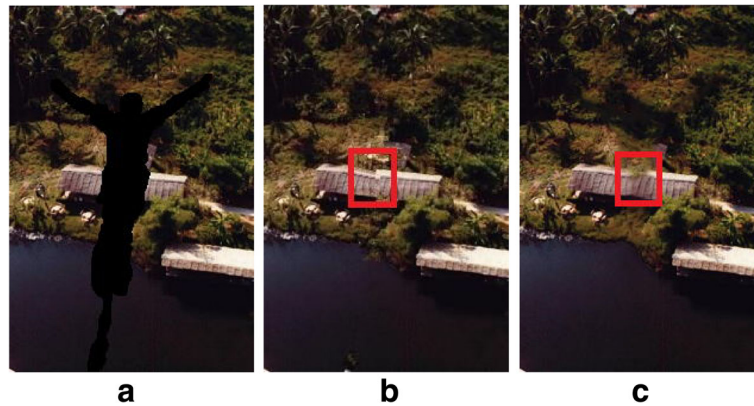


Fig. 11 A comparison of inpainted results using fixed patch and dynamic patch. **a** Mask. **b** A fixed patch. **c** Dynamic patches

works, we consider that our method can also be applied to image compression field. So in the future, we will expand our work to image compression task. And we expect that our approach can lead to the satisfactory visual effect and higher compression ratio.

5 Comparison between the proposed method and deep learning

Recently, deep learning has been diffusely applied to image completion to extract high-level features of images. Pathak et al. [37] proposed context encoders motivated by feature learning which used a convolutional neural network (CNN) trained with a reconstruction plus an adversarial loss to generate the contents of an arbitrary image region. Iizuka et al. [38] proposed a novel architecture based on context encoders with global and local context discriminators, which resulted in both globally and locally consistent images. Yu et al. [39] presented a unified feed-forward generative network using a novel contextual attention layer for image inpainting, which is trained end to end with reconstruction losses and two Wasserstein GAN losses to generate higher-quality inpainting results.

Undoubtedly, deep learning-based inpainting methods may achieve more meticulous results by utilizing a convolutional neural network to extract high-level features. However, image inpainting algorithms using the high-level features require very complex models and a large amount of parameters to adjust, especially the higher the accuracy

of the model is, the worse the general robustness will be. And they need also a number of external training data that needs to be collected in the same circumstances as the data set used. In a fast-developing society, the quality of the inpainting result is not the only criterion to evaluate the effectiveness of the algorithm, since the runtime is also a significant factor that cannot be ignored. Therefore, the proposed method is more robust for image inpainting based on low-level features. And our method has fewer parameters and can be adjusted easily. Furthermore, the computational time is much shorter than that of deep-learning methods. Our method will be further improved in the future to achieve better visual quality.

6 Conclusions

In this paper, we have proposed a novel inpainting method which estimates the missing pixels by MLS method in 3D subspace as the prior knowledge utilized to solve the confidence term. And we also add a curvature factor for the data term to avoid its value of zero. We propose an angle-aware rotation patch matching strategy which considers the different angles of the same source patch in order to find multiple candidate patches for every target patch, thereby increasing the matching accuracy. Meanwhile, a Jaccard similarity coefficient is used to enhance the textural and structural similarity between the target patch and each corresponding source patch. In order to improve the restored efficiency, we divide the whole image into high-frequency and low-frequency components and introduce the gradient to determine the size of target patch dynamically. Our method is applied to the filling with a large hole, the restoration of multiple missing blocks and object removal task. And a large number of experimental results demonstrate that the proposed method is superior to other advanced methods and the runtime is shorter. In the future, we will extend our approach to image compression field.

Table 1 Comparison of mean run time in HH:MM:SS for 30 images

Algorithms	Runtime (average)
Criminisi	00:02:10
GSR	00:10:16
Melding	00:08:57
Aloha	02:24:36
Ours	00:03:45

Abbreviations

ALOHA: Annihilating filter-based low-rank Hankel matrix approach; BP: Back propagation; CCD: Curvature-driven diffusions model; CNN: Convolutional neural network; EM: Expectation maximization; GAN: Generative adversarial networks; GSR: Group-based sparse representation; MLS: Moving least squares; MRF: Markov random field; NNF: Nearest neighbor field; PSNR: Peak signal-to-noise ratio; SSD: Sum of squared differences; SSIM: Structure similarity index; TV: Total variation

Acknowledgements

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

Authors' contributions

All authors take part in the discussion of the work described in this paper. All authors read and approved the final manuscript.

Funding

This paper is supported by the National Natural Science Foundation of China (grant nos. 61772322, 61572298, 61702310, and 61873151).

Availability of data and materials

Please contact authors for data requests.

Competing interests

The authors declare that they have no competing interests.

Received: 25 February 2019 Accepted: 13 June 2019

Published online: 08 July 2019

References

- M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, *Image inpainting* (Proceedings of ACM SIGGRAPH, ACM Press, 2000), pp. 417–424
- J. Shen, T.F. Chan, Mathematical models for local nontexture inpaintings. *SIAM Journal on Applied Mathematics* **62**, 1019–1043 (2002)
- T.F. Chan, J. Shen, Non-texture inpainting by curvature-driven diffusion (CCD). *J. Vis. Commun. Image Represent* **12**, 436–449 (2001)
- Z. Xu, X. Lian, L. Feng, Image in-painting algorithm based on partial differential equation. *ISECS Intl. Colloq. Comput. Commun. Control Manag* **1**, 120–124 (2008)
- Y.W. Wen, R.H. Chan, A.M. Yip, A primal-dual method for total variation based wavelet domain inpainting. *IEEE Trans. Image Process.* **21**(1), 106–114 (2012)
- R.L. Biradar, V.V. Kohir, A novel image inpainting technique based on median diffusion. *Sadhana* **38**(4), 621–644 (2013)
- Q. Cheng, H. Shen, L. Zhang, P. Li, Inpainting for remotely sensed images with a multichannel nonlocal total variation model. *IEEE Trans. Geosci. Remote Sens.* **52**(1), 175–187 (2014)
- V.B.S. Prasath, D.N.H. Thanh, H.H. Hai, N.X. Cuong, *Image restoration with total variation and iterative regularization parameter estimation. The 8th International Symposium on Information and Communication Technology (SoICT 2017)* (ACM, 2017)
- A.A. Efros, T.K. Leung, in *ICCV*. Texture synthesis by non-parametric sampling (1999)
- A. Criminisi, P. Perez, K. Toyama, *Object removal by exemplar-based inpainting. Inter.conf.computer Vision and Pattern Recog Cvpr*, vol 2 (2003), pp. 721–728
- J. Sun, L. Yuan, J. Jia, H.Y. Shum, Image completion with structure propagation. *ACM Trans. Graph.* **24**(3), 861–868 (2005)
- A. Wong, J. Orchard, A nonlocal-means approach to exemplar-based inpainting. *IEEE International Conference on Image Processing. IEEE* (2008)
- D. Ding, S. Ram, J. Rodriguez, Perceptually aware image inpainting. *Pattern Recognition*. **83**, 174–184 (2018)
- S. Li, M. Zhao, Image inpainting with salient structure completion and texture propagation. *Pattern Recognition Letters*. **32**(9), 1256–1266 (2011)
- J. Wang, K. Lu, D. Pan, N. He, B.-K. Bao, Robust object removal with an exemplar-based image inpainting approach. *Neurocomputing* **123**, 150–155 (2014)
- B.S. Kim, J.S. Kim, J. Park, Exemplar based inpainting in a multi-scaled space. *Optik - International Journal for Light and Electron Optics*. **126**(23), 3978–3981 (2015)
- B. Liu, P. Li, B. Sheng, Y.W. Nie, E.H. Wu, Structure-preserving image completion with multi-level dynamic patches. *Visual Computer*. **8**, 1–14 (2017)
- H. Wang, Y. Cai, R. Liang, X.X. Li, Exemplar-based image inpainting using structure consistent patch matching. *Neurocomputing*. **269**, 401–410 (2017)
- N. Komodakis, G. Tziritis, Image completion using efficient belief propagation via priority scheduling and dynamic pruning. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*. **16**(11), 2649–2661 (2007)
- T. Ruzic, A. Pizurica, Context-aware patch-based image inpainting using Markov random field modeling. *IEEE Transactions on Image Processing*. **24**(1), 444–456 (2015)
- N. Alotaibi, *Image completion by structure reconstruction and texture synthesis* (Springer-Verlag, 2015)
- S. Ge, K. Xie, R. Yang, Z.Q. Shi, Image completion using global patch matching and optimal seam synthesis. *International Conference on Pattern Recognition. IEEE Computer Society* (2014)
- H. L. Zhao, H. Y. Guo, X. G. Jin, J. B. Shen, X. Y. Mao, J. R. Liu, Parallel and efficient approximate nearest patch matching for image editing applications. *Neurocomputing*. S0925231218304703 (2018)
- J.B. Huang, S.B. Kang, N. Ahuja, J. Kopf, Image completion using planar structure guidance. *Acm Transactions on Graphics*. **33**(4), 1–10 (2014)
- Q. H. Zeng, L. U. De-Tang, Curve and Surface Fitting Based on Moving Least-Squares Methods. *Journal of Engineering Graphics*. **25**(1), 84–89 (2004)
- L. Yin, C. Chang, An effective exemplar-based image inpainting method. *IEEE, International Conference on Communication Technology*. 739–743 (2012)
- K. Li, Y. Wei, Z. Yang, W.H. Wei, Image inpainting algorithm based on TV model and evolutionary algorithm. *Soft Computing*. **20**(3), 885–893 (2016)
- S. Darabi, E. Shechtman, C. Barnes, D.B. Goldman, P. Sen, Image melding: combining inconsistent images using patch-based synthesis. *ACM Transactions on Graphics*. **31**(4) (2012)
- J. Zhang, D. Zhao, W. Gao, Group-based sparse representation for image restoration. *IEEE Transactions on Image Processing*. **23**(8), 3336–3351 (2014)
- K.H. Jin, J.C. Ye, Annihilating filter-based low-rank Hankel matrix approach for image inpainting. *IEEE Transactions on Image Processing*. **24**(11), 3498–3511 (2015)
- M.P. Eckert, A.P. Bradley, Perceptual quality metrics applied to still image compression. *Signal Processing* **70**(3), 177–200 (1998) 32. Z. Wang, A. C. Bovik, H. R. Sheikh, E.P. Simoncelli
- Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*. **13**(4), 600–612 (2004)
- D. Liu, X. Y. Sun, F. Wu, S. P. Li, Y.-Q. Zhang, Image Compression With Edge-Based Inpainting. *IEEE Transactions on Circuits and Systems for Video Technology*, **17**(10), 1273–1287 (2007)
- C. Qin, C. C. Chang, Y.-P. Chiu, A Novel Joint Data-Hiding and Compression Scheme Based on SMVQ and Image Inpainting. *IEEE Transactions on Image Processing*. **23**(3), 969–978 (2014)
- C. Qin, Q. Zhou, F. Cao, J. Dong, X. P. Zhang, Flexible Lossy Compression for Selective Encrypted Image with Image Inpainting. *IEEE Transactions on Circuits and Systems for Video Technology*, 1–1 (2018)
- B. Coloma, B. Marcelo, C. Vicent, G. Luis, M. AdriSn, R. Florent, An Inpainting-Based Deinterlacing Method. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*. **16**(10), 2476–91 (2007)
- D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, Context encoders: Feature learning by inpainting. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2536–2544 (2016)
- S. Iizuka, E. Simo-Serra, H. Ishikawa, Globally and locally consistent image completion. *ACM Transactions on Graphics (ToG)*, **36**(4), 107 (2017)
- J. H. Yu, Z. Lin, J. M. Yang, X. H. Shen, Generative image inpainting with contextual attention. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5505–5514 (2018)

7 Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.