A Novel Image Completion Algorithm Based on Planar Features

Mang Xiao[1*], Yunxiang Liu[1], Li Xie[2], Qiaochuan Chen[2], Guangyao Li[2]
[1] School of Computer Science and Information Engineering,
Shanghai Institute of Technology, Shanghai 201418, China
[e-mail: mangxiao@sit.edu.cn]
[2] College of Electronics and Information Engineering,
Tongji University, Shanghai, China
*Corresponding author: Mang Xiao

Received November 18, 2016; revised March 28, 2017; revised December 8, 2017; accepted February 22, 2018; published August 31, 2018

Abstract

A novel image completion method is proposed that uses the advantage of planar structural information to fill corrupted portions of an image. First, in estimating parameters of the projection plane, the image is divided into several planes, and their planar structural information is analyzed. Second, in calculating the a priori probability of patch and patch offset regularity, this information is converted into a constraint condition to guide the process of filling the hole. Experimental results show that the proposed algorithm is fast and effective, and ensures the structure continuity of the damaged region and smoothness of the texture.

Keywords: Image completion, image inpainting, planar features, global optimization, guided synthesis

This work was supported by National Nature Science Foundation of China (NSFC) (60771065, 51378365), Foundation of Shanghai institute of Technology (YJ2017-5) and Technology Projects of education bureau of Jiangxi province of China (GJJ151001).

http://doi.org/10.3837/tiis.2018.08.016 **ISSN : 1976-7277**

1. Introduction

In the revolution of intelligent technologies, massive digital images continue to emerge. However, many factors can cause digital image defects [1] [2]. On the one hand, digital equipment collecting the raw data cannot meet the actual needs of diversification. Thus, we need to edit the image, repair defects (e.g., scratches) appearing in the image or removing a distinct object from an image, and so on. On the other hand, the image is processed by acquisition, encoding, storage, transmission, and post-processing, all of which tend to create a certain degree of distortion, seriously impacting image quality. Therefore, image completion is vital to ensuring image integrity and enhancing image aesthetics.

A digital image completion technique is the use of the information of a known region in an image to fill an unknown region (a hole) in accordance with certain rules. After image completion, it is possible for the result to achieve a natural visual image for the human eye, which will not detect traces of the repair. As the research hotspot of computer graphics and vision, image completion technology is widely used in the protection of cultural relics, film special effects, image compression, image analysis of ultra-high resolution, video transmission error concealment, and other image processing applications.

Depending on the principle used, an image restoration algorithm is roughly divided into two categories [2]: diffusion-based and exemplar-based. The former is primarily used for small-scale damaged regions, as such methods contain model partial differential equations and a variational model [3] [4].

The latter is primarily used for large damaged regions, including synthetic texture-based sample blocks [5] [6] [7] [8], based on a transform domain such as Fourier transform and wavelet transform [9]. Such methods usually adopt the image exemplar (patch) as the basic unit of operation. First, according to information at the boundary of the unknown region, the best matching patches are searched from known areas of the image through the relevant calculations. Finally, the best matching patches are used to fill the damaged region to complete the image. However, the above methods have difficulty in maintaining the structural continuity of the image.

In recent years, using a random search algorithm known as patch match, Barnes et al. [10] proposed an approximate solution to quickly locate a similar patch, disseminate the location information to the neighboring patches, and then find the most similar patch for enhancing image structural coherence. Subsequently, the patch is rotated and scaled. Later, Barnes et al. [11] expanded the patch match algorithm. Similarity measurements between patches use colors from the corresponding pixels in addition to feature descriptors, enhancing the robustness of the algorithm. As patch match is effectively limited in patch search space, and reducing noise or using an irrelevant search could result in a local minimum and improve the accuracy of a matching patch. However, for higher resolution images, the complexity of this method is high. He et al. [12] built a k-dimensional tree (kd-tree) to search for the best similar patch, and greatly improved the running time over [10].

However, a larger storage space is required, making it difficult to extend the general transformation (such as rotation and zoom). Huang et al. [13] adopted the dividing plane and extraction plane regularity to extend the patch match, greatly enhancing the effect of image completion. Li et al. [21] proposed a joint multiplanar autoregressive and low-rank based approach for image completion. This method outperforms on the high pixel missing rate. While for the huge removed region, this approach needs much time to run and is difficult to obtain the satisfactory results. So they proposed another method [22] to resolve this problem and gain better results using structure-guided and regularity statistics. Moreover, using structure propagation and structure-guided completion, Baek et al. [23] proposed a multiple image completion method that preserves the geometric consistency among different views.

We observe that most of these algorithms have two important limitations. The first one is that can't find the best source patch for the target patch, which especial in the image including many planes. The other one is hard to construct the patch formation for the scenes that are not fronto-parallel. These two limitations seriously affect the effect of the image completion. Therefore, this study presents a new approach that utilizes planar structural information to guide the image completion, which can fully utilize the image information of the texture and structure in addition to completing the image in different scenarios.

The primary work includes three main elements. The first one is improving the similar measure method for the patches, adding a new gradient term to measure the similarity between patches. The next one expands the patch searching space to different planes, enhancing the accuracy of the similar patches. The last one is using the plane regularity to construct the patch transformation, which can address the difficult situation of image completion in different planes.

2. Image Completion

The proposed image completion algorithm includes the following four steps. The framework is shown in **Fig. 1**.

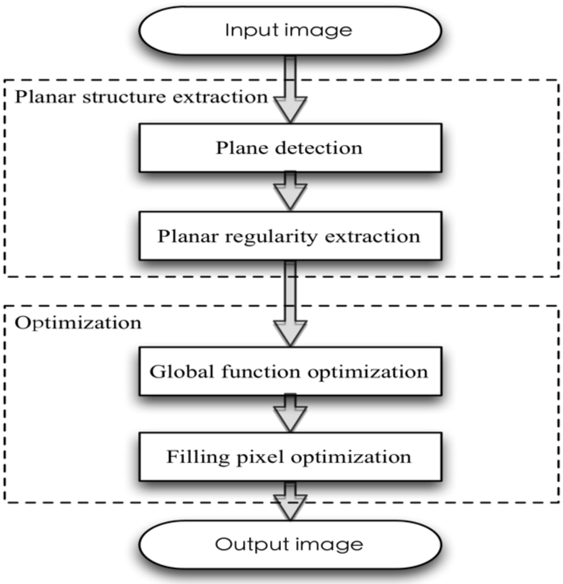

Fig. 1. The framework of our method.

2.1 Planar structure extraction

This section describes the extraction of planar structural information from a known image area, including plane detection and planar regularity extraction.

2.1.1 Plane detection

In recent years, many efficient plane detection algorithms [14] have been proposed, which can transform a distorted perspective plane into a positive parallel plane. This process includes three steps: line extraction, vanishing point estimation, and plane orient detection. A brief overview follows, as this part of the algorithm is quite standard.

First of all, line extraction is performed in the known region in the image. Then based on the same random sampling method of voting, vanishing points are estimated for different planes. Finally, the plane orientation is computed using the vanishing points by which two different orient point lines can define a unique plane.

2.1.2 Planar regularity extraction

In many artificial environment and architecture scenes, there often exists some structure regularity. This regularity detection has proven to be a simple and direct means of understanding the structure of a scene. In order to extract the planar regularity of each plane, we adopt a detection algorithm to correct the affine or distorted plane.

First, the Gaussian feature points algorithm difference is detected in the known region of the image. We use the Gaussian feature points algorithm with reference to Huang [13]. Because an affine correction would seriously distort the source surface of the image, we choose the original image plane rather than the correction plane. Moreover, the scale-invariant feature transform (SIFT) [15] is calculated and a kd-tree is adopted to calculate the nearest neighbors for each patch. Third, for each plane, all features matching two features having a higher position in an a priori probability are extracted. Finally, the affine correction feature point position is matched to correct the perspective distortion of repeating structures in two-dimensional space. This includes the displacement of the two correction feature points over the same space, such that we can detect a duplicate translation. If a structure law exists, this will be the two-dimensional displacement vector space affine correction for a dense cluster. Using the mean shift algorithm [16], we can compute these patterns.

2.2 Planar structure guided energy function

2.2.1 Energy function optimization

The planar structure guided global energy function is a measured distance function that includes three main elements. The closer the distance, the greater the correlation, and the higher the similarity. There are two ways to enhance our global objective function of image completion. In order to improve the objective function, a gradient term is added into the color elements of this study. Then the search space is determined by the plane index to determine the patch mapping scheme.

The planar structure guided energy function is defined as follows:

$$\min_{\{t_i,s_i,m_i\}} \sum_{i\in\bar{\Omega}} E_{color}(s_i,t_i,m_i) + E_{guide}(s_i,t_i,m_i) + E_{gradient}(\nabla s_i,\nabla t_i,m_i) \tag{1}$$

where E_{color} is the color item, E_{guide} is the planar structure guide item, $E_{gradient}$ is the gradient item, Ω is the known region, $\bar{\Omega}$ is the damaged region, $t_i = (t_i^x, t_i^y)^T$ is the i-th target patch that is centered in the pixel at (x, y), $s_i = (s_i^x, s_i^y)^T$ is the i-th source patch, and m_i is the plane index of target patch t_i.

The color item is defined as follows:

$$E_{\text{color}}(s_i, t_i, m_i) = \left\| q(s_i, t_i, m_i) - p(t_i) \right\|^2 \quad (2)$$

where $p(t_i)$ are the pixels of a 7×7 t_i, and $q(s_i, t_i, m_i)$ are the pixels of a 7×7 s_i that corresponds to t_i from the plane m_i.

The planar structure guide item is defined as follows:

$$E_{guide}(s_i, t_i, m_i) = \lambda_1 E_{plane}(s_i, t_i, m_i) + \lambda_2 E_{direction}(s_i, t_i, m_i) + \lambda_3 E_{proximity}(s_i, t_i) \quad (3)$$

where E_{plane} is the cost of the plane to which source patch s_i belongs, $E_{direction}$ is the cost of the plane direction to which source patch s_i belongs, and $E_{proximity}$ is the cost of the similar proximity between source patch s_i and target patch t_i. As a matter of experience, we set $\lambda_1 = 10, \lambda_2 = 10^3, \lambda_3 = 1$, respectively.

The plane item is defined as follows:

$$E_{plane}(s_i, t_i, m_i) = -\log \Pr\left[m_i \mid s_i \right] - \log \Pr\left[m_i \mid t_i \right] \quad (4)$$

where the $\Pr\left[m_i \mid s_i \right]$ is the probability of source patch s_i belonging to plane m_i, and $\Pr\left[m_i \mid t_i \right]$ is the probability of target patch t_i belonging to plane m_i. In this study, the negative logarithm of the likelihood function is used to convert the plane probability into a penalty term. This supports that the plane has a high probability of sampling in the source patch and target patch.

The plane direction item is defined as follows:

$$E_{direction}(s_i, t_i, m_i) = \psi\left(\min\left(\left| H_{m_i}^1(\tilde{s}_i)^y - H_{m_i}^1(\tilde{t}_i)^y \right|, \left| H_{m_i}^2(\tilde{s}_i)^y - H_{m_i}^2(\tilde{t}_i)^y \right| \right) \right) \quad (5)$$

where $\psi(z) = \min(|z|, c)$ is the limitation function, and the value of c is set to 0.02. We define target patch $\tilde{t}_i = \left[t_i^x, t_i^y, 1 \right]^T$, and source patch $\tilde{s}_i = \left[s_i^x, s_i^y, 1 \right]^T$. Corresponding in the affine correction space, the target patch is $\tilde{t}_i' = \left[h_1 \tilde{t}_i, h_2 \tilde{t}_i, h_3 \tilde{t}_i \right]^T$, and source patch is $\tilde{s}_i' = \left[h_1 \tilde{s}_i, h_2 \tilde{s}_i, h_3 \tilde{s}_i \right]^T$, where h_1, h_2, h_3 are row vectors of the perspective transformation matrix H_m. H_m, defined as follows:

$$H_m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1^m & l_2^m & l_3^m \end{bmatrix} \tag{6}$$

where $l_\infty^m = \left[l_1^m, l_2^m, l_3^m \right]^T$ denote the vanishing line, which can represent a plane with two lines of different direction.

The offset (d^x, d^y) is defined as the correct space displacement vector from the target patch $\tilde{t}_i$ to the source patch $\tilde{s}'_i$ in the affine correction space. Thus, $\tilde{s}'_i$ can be computed as follows:

$$\tilde{s}'_i = \begin{bmatrix} h_1 + h_3 d^x \\ h_1 + h_3 d^y \\ h_3 \end{bmatrix} \tilde{t}_i \tag{7}$$

Source patch $\tilde{s}_i$ is computed using inverse matrix H_m^{-1} of perspective transformation matrix H_m as follows:

$$\tilde{s}_i = H_{m_i}^{-1} \tilde{s}'_i = H_{m_i}^{-1} \begin{bmatrix} h_1 + h_3 d^x \\ h_2 + h_3 d^y \\ h_3 \end{bmatrix} \tilde{t}_i \tag{8}$$

In order to obtain the motion parameters of patch s_i, we apply an offset of the transformation matrix:

$$\tilde{s}_i = H_{m_i}^{-1} \begin{bmatrix} h_1 + h_3 d^x \\ h_2 + h_3 d^y \\ h_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & t_i^x \\ 0 & 1 & t_i^y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = T_{s_i} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{9}$$

Generally, a building image consists of horizontal and vertical directions of repeating structures. This encourages locating an orthogonal source patch going in either direction. It is noteworthy that affine correction tolerates each parallel to the vanishing point of support. Mapping the vanishing point of each segment set, and aligning with the horizontal axis, the perspective transformation matrix of each plane is calculated using the following equation:

$$H_{m_i}^j = \begin{bmatrix} \cos(\theta_j) & -\sin(\theta_j) & 0 \\ \sin(\theta_j) & \cos(\theta_j) & 0 \\ l_1^{m_i} & l_2^{m_i} & l_3^{m_i} \end{bmatrix} \tag{10}$$

In addition to the restrictions given above, this study further introduces fewer restrictions, which limits copying matching blocks from a dramatically different region. The proximity item is defined as follows:

$$E_{proximity}(s_i,t_i)=\frac{\left\|s_i-t_i\right\|_2^2}{\sigma_d(t_i)^2+\sigma_c^2} \tag{11}$$

where $\sigma_d(t_i)^2$ is the square of the distance from the nearest known pixels to the target area boundary. $\sigma_c^2=\left(W/8\right)^2$ as the intensity adjustment parameter.

In recent years, Arias et al. [17] applied the gradient of the patch for image completion, and used the L1 norm to measure the gradient term, which can enhance the effect of the patch matching algorithm in dealing with a high detail texture region. Darabi et al. [18] also introduced a gradient term into the optimization function and used the L0 norm to prevent the loss of image detail when mixing gradients and increase the search space patch matching to achieve the texture deformation from a different texture image. Given that the gradient is significant for image completion, it is also introduced into our global energy function:

$$E_{gradient}(\nabla s_i,\nabla t_i,m_i)=\left\|\nabla q(s_i,t_i,m_i)-\nabla p(t_i)\right\|^2 \tag{12}$$

where $\nabla p(t_i)$ is the gradient of target patch t_i, and $\nabla q(s_i,t_i,m_i)$ is the gradient of source patch s_i. For image completion based on patch matching, when the source and target patches are slightly different in color and texture, a gradient term can play to our strengths and search for the best similar patch for higher consistency.

2.2.2 Filling pixel optimization

(1) Searching for the best similar patch

To speed up the search for the best similar patch, we use the patch match method, which finds the corresponding best matching patch in the source image for the target patch. When searching, the location of a matching patch is always found first. Then, we search for a matching transformed patch. Thus, the energy function value is minimized.

When updating the filling pixel, we calculate the correlation distance between the pixels that are associated with a higher degree to the source image pixels near the area boundary defect, by weight $w(x,y)=2^{-d(x,y)}$, where $d(x,y)$ represents the shortest distance of the pixel to the boundary.

This is because after the planar transformation, the sample patch of space expands much more than the original, and as such, requires more iterations to converge to an optimal value. Biasing the random sample patch mapping algorithm to generate a random sample patch area is more likely to improve the convergence rate. Meanwhile, the transformation parameters additionally set the corresponding minimum and maximum thresholds. Thus, increasing the algorithm spread does not affect the accuracy of the optimal value.

(2) Filling pixel

Wexler et al. used a weighted voting method to calculate the pixel value that achieved good visual results. In this study, a similar calculation method is adopted to obtain good visual results of image completion.

3. Experimental results and analysis

In order to verify the validity of the proposed method, we used an experimental platform with an Intel (R) Core (TM) i7-4770K CPU, 3.5GHz, 16GB of RAM, and Matlab R2015a. From the scientific and objective consideration of the experiment, we used the visual and peak signal to noise ratio (PSNR) evaluation methods to analyze the experimental results.

3.1 Visual evaluation method

As can be seen in the following figures, the first column displays the input images for filling, the second column displays the results using Photoshop, the third column displays the results using the Huang et al. [13] algorithm, the fourth column displays the results using our method, the fifth column displays the results utilizing the Darabi et al. [18] approach. For **Figs. 2** and **3**, the algorithms of Photoshop and Huang performed poorly, which can be observed in the discontinuity of the building structures. However, the image results of our method display continuity and harmony. For **Figs. 4** and **5**, the Photoshop, Darabi, and Huang methods did not repair the handrail structure or maintain image continuity. The proposed algorithm can maintain the structural information of the image well and can achieve good visual effect.

Fig. 2. Skewed building.

Fig. 3. Railing of the park.

Fig. 4. Streets and building.

Fig. 5. Stairs.

Fig. 6. Square and buildings.

Fig. 7. Steel bridge.

Example Results: The first column shows an unknown region (white). Columns 2 through 4 are the results from Photoshop, Huang, our algorithm, and Darabi, respectively.

3.2 PSNR evaluation method

In this paper, we adopted a PSNR method to objectively evaluate the outcome of the experimental results. A larger PSNR value denotes that the completed image has the higher performing algorithm. The image completion performance of the **Figs. 2** through **7** is shown in **Table 1**.

Table 1. Image completion performance evaluated in PSNR.

Image No.	Photoshop	Huang	Darabi	Proposed
1	24.36	24.84	24.59	25.17
2	26.52	26.43	26.3	27.09
3	23.15	23.34	23.46	23.85
4	24.28	24.73	24.61	25.46
5	24.12	24.46	24.32	25.05
6	25.61	25.94	25.79	26.38
mean	24.673	24.957	24.845	25. 5

It is obvious that the PSNR of our method is the highest on each image. The PSNR of Photoshop is almost the lowest on each image. As shown in **Table 2**, the average PSNR image value of the Photoshop method is 24.673, the Huang method is 24.958, the Darabi algorithm is 24.845, and our method, with the highest, is 25.5. For each image, the PSNR of our method is the highest, which can be easily determined from **Fig. 8**. Therefore, our method has the highest performance demonstrated with the PSNR on filling the damaged portions of the image.

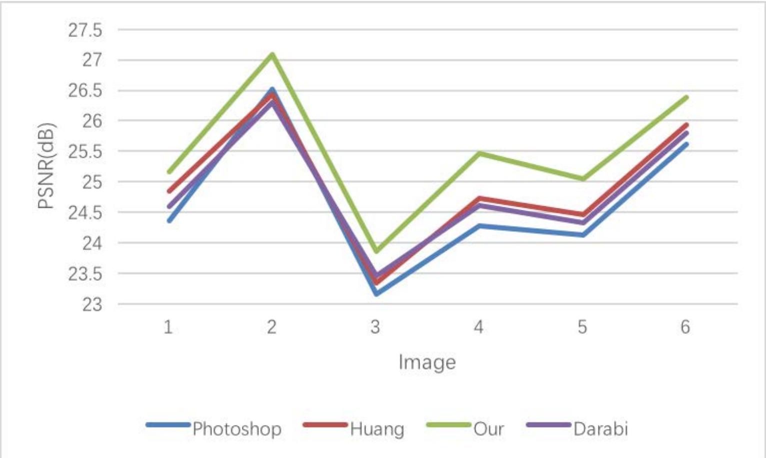

Fig. 8. PSNR comparison.

3.3 SSIM evaluation method

In order to further verify the superiority of our method, structural similarity (SSIM) [19], the third method of image quality evaluation is adopted to evaluate the effectiveness of the image completion results. According to the theory of structural similarity, the image signal has a high degree of structure. Moreover, the pixels are closely related, especially the pixels belonging to the same domain, which reflect the structural information of the real scene.

The SSIM value reflects the similarity between two images. When the SSIM value is larger, the two images are more alike. On the other hand, when the SSIM value is smaller, the similarity between the two images is lower. When performing an SSIM analysis, the images in **Fig.s 2** through **7** can be divided into two categories. **Fig. 2** and **Fig. 4** comprise the first category, where the deleted contents of the target area are the real parts of the scene's structure. Other figures comprise the second category, where the removed target area is another object, such as a person, and not an actual part of the scene's structure. In general, the removed content usually differs from the color and structure of the scene image. Therefore, in most cases, the SSIM value of the second category is smaller than the value of the first.

As shown in **Table 2**, the average SSIM image value of the Photoshop method is 0.878, the Huang method is 0.885, the Darabi algorithm is 0.884, and our method, with the highest, is 0.894. For each image, the SSIM of our method is the highest, which can be easily determined from **Fig. 9**. These comparisons show that the image completion results of our method can maintain better structural similarity.

Table 2. Image completion performance evaluated in SSIM.

Image No.	Photoshop	Huang	Darabi	Proposed
1	0.922	0.928	0.926	0.937
2	0.781	0.792	0.789	0.816
3	0.941	0.947	0.948	0.951
4	0.752	0.759	0.761	0.768
5	0.929	0.934	0.931	0.938
6	0.945	0.951	0.949	0.954
mean	0.878	0.885	0.884	0.894

Through human vision, it is easy to discern that the image completion results of **Figs. 2** through **7** using our method are better than the results using the other two methods. Our approach is better at preserving image structure continuity and consistency, making the image completion result consistent with the human visual system.

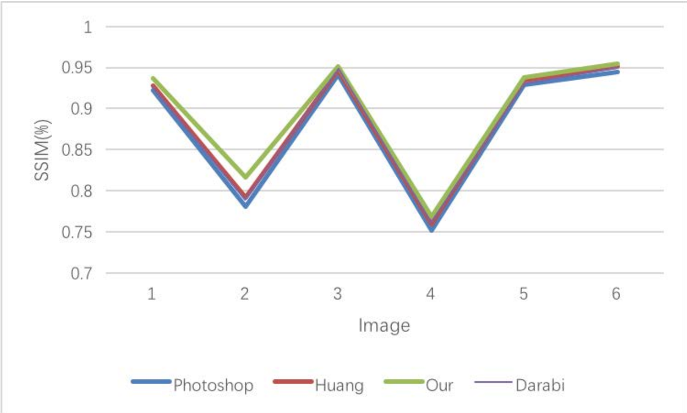

Fig. 9. SSIM comparison.

Generally, the PSNR of our method is slightly higher than the PSNR of other algorithms, as shown in **Fig. 6**. Similarly, the SSIM values of our method are generally slightly higher than SSIM values of other algorithms. Therefore, through the human vision system, PSNR and SSIM value comparisons show that our algorithm is more superior.

4. Conclusions

For building the image plane structure, there is often repetitive regularity. This paper proposes an image texture, without human intervention and structural information, by estimating planar projection parameters analyzed between the translation law planes converted into a soft constraint to guide the approximate block matching method.

This paper presents a novel image completion method that uses the advantage of planar structural information to fill damaged portions of the image. First, in estimating parameters of the projection plane, the image is divided into several planes, and the planar structural information is analyzed. Furthermore, in calculating the a priori probability of the patch and patch offset regularity, this information is converted into a constraint condition to guide the process of filling the hole.

The proposed algorithm possesses a high degree of robustness and can detect both a plurality of planes and corresponding patch offset regularity of the image, and handle different types of images to be repaired. Furthermore, the proposed image processing method is fast and effective, and ensures structural continuity of the damaged area and textural smoothness.

References

[1] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester, "Image inpainting," in *Proc. of Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, New Orleans, Louisiana, USA, pp. 417-424, 2000. Article (CrossRef Link)

[2] C. Guillemot, O. Le Meur, "Image inpainting: Overview and recent advances," *Signal Processing Magazine, IEEE*, vol. 31, no. 1, pp. 127-144, 2014. Article (CrossRef Link)

[3] T. F. Chan, J. Shen, "Nontexture inpainting by curvature-driven diffusions," *Journal of Visual Communication and Image Representation*, vol. 12, no. 4, pp. 436-449, 2001. Article (CrossRef Link)

[4] M. Bertalmio, L. Vese, G. Sapiro, S. Osher, "Simultaneous structure and texture image inpainting," *Image Processing, IEEE Transactions on*, vol. 12 no. 8, pp. 882-889, 2003. Article (CrossRef Link)

[5] A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proc. of IEEE International Conference on Computer Vision*, Greece, pp. 1033-1038, 1999. Article (CrossRef Link)

[6] M. Xiao, G. Y. Li, L. Xie, Y. L. Tan, Y. H. Mao, "Contour-guided image completion using a sample image," *Journal of Electronic Imaging*, vol. 24, no. 2, pp. 023029-023029, 2015. Article (CrossRef Link)

[7] P. P. Criminisi and K. Toyama, "Region filling and object removal by exemplar-based image in painting," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200-1212, 2004. Article (CrossRef Link)

[8] M. Xiao, G. Y. Li, Y. Y. Jiang, L. Xie, Y. He, "Image Completion using Belief Propagation Based on Planar Priorities," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 9, pp. 4405-4418, 2016. Article (CrossRef Link)

[9] U. A. Ignácio, C. R. Jung, "Block-based image inpainting in the wavelet domain," *The Visual Computer*, vol. 23, no. 9, pp. 733-741, 2007. Article (CrossRef Link)

[10] C. Barnes, E. Shechtman, A. Finkelstein, D. Goldman, "PatchMatch: A randomized correspondence algorithm for structural image editing," *ACM Transactions on Graphics-TOG*, vol. 28, no. 3, pp. 24, 2009. Article (CrossRef Link)

[11] C. Barnes, E. Shechtman, D. B. Goldman, A. Finkelstein, "The generalized patchmatch correspondence algorithm," *Computer Vision–ECCV*, Springer Berlin Heidelberg, pp. 29-43, 2010. Article (CrossRef Link)

[12] K. He, J. Sun, "Computing nearest-neighbor fields via propagation-assisted kd-trees," in *Proc. of Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, pp. 111-118, 2012. Article (CrossRef Link)

[13] J. B. Huang, S. B. Kang, N. Ahuja, "Image completion using planar structure guidance," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 4, pp. 129, 2014. Article (CrossRef Link)

[14] O. Chum, J. Matas, "Planar affine rectification from change of scale," *Computer Vision–ACCV*, Springer Berlin Heidelberg. pp. 347-360, 2011. Article (CrossRef Link)

[15] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91-110, 2004. Article (CrossRef Link)

[16] D. Comaniciu, P. Meer, "Mean shift: A robust approach toward feature space analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 5, pp. 603-619, 2002. Article (CrossRef Link)

[17] P. Arias, G. Facciolo, V. Caselles, G. Sapiro, "A variational framework for exemplar-based image inpainting," *International journal of computer vision*, vol. 93, no. 3, pp. 319-347, 2011. Article (CrossRef Link)

[18] S. Darabi, E. Shechtman, C. Barnes, D. B. Goldman, P. Sen, "Image melding: combining inconsistent images using patch-based synthesis," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 82, 2012. Article (CrossRef Link)

[19] Z. Wang, A. C. Bovik, H. R. Sheikh, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol.13 no. 4, pp. 600-612,2004. Article (CrossRef Link)

[20] A. Newson, A. Almansa, M. Fradet, Y. Gousseau, P. Prez, "Video inpainting of complex scenes," SIAM Journal on Imaging Sciences, vol. 7, no. 4, pp. 1993-2019, 2014. Article (CrossRef Link)

[21] S. Yang, J. Liu, S. Song, M. Li, Z. Guo, "Structure-guided image completion via regularity statistics," in *Proc. of Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on. IEEE*, pp. 1711-1715, 2016. Article (CrossRef Link)

[22] M. Li, J. Liu, Z. Xiong, X. Sun, Z. Guo, "MARLow: A Joint Multiplanar Autoregressive and Low-Rank Approach for Image Completion," in *Proc. of European Conference on Computer Vision*, Springer International Publishing, pp. 819-834, 2016. Article (CrossRef Link)

[23] S. H. Baek, I. Choi, M. H. Kim, "Multiview image completion with space structure propagation," in *Proc. of Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 488-496, 2016. Article (CrossRef Link)

Mang Xiao received his BS degree from Jiangxi Normal University in 2005, and his MS degree in computer science from Nanchang University in 2010. He received his PhD from Tongji University in 2016. Now he is a lecturer in the School of Computer Science and Information Engineering, Shanghai Institute of Technology. His research interests include image processing and machine learning.

Yunxiang Liu received his BS and MS degrees from Northeast Normal University, Jinlin, in 1986 and 1989, respectively. He received his PhD from Jinlin University in 1997. Now he is a professor and master supervisor in the School of Computer Science and Information Engineering, Shanghai Institute of Technology. His primary research interests include Artificial Intelligence, Computer Software and Theory.

Li Xie received his BS degree from Zhejiang University of Technology in 2012. Now, he is a PhD candidate in the College of Electronics and Information Engineering at Tongji University. His research interests include image processing and data mining.

Qiaochuan Chen received his BS degree from Hefei University of Technology in 2013, and his MS degree in computer science from Yunnan University in 2015. Now, he is a PhD candidate in the College of Electronics and Information Engineering at Tongji University. His research interests include image processing and data mining.

Guangyao Li received his BS and MS degrees from Nanjing University of Aeronautics and Astronautics, Nanjing, in 1986 and 1989, respectively. He received his PhD from Nanjing University of Aeronautics and Astronautics in 1997. Now he is a professor and doctoral supervisor in the College of Electrical and Information Engineering, Tongji University. His primary research interests include image processing and virtual reality.