SUPERPIXEL-BASED IMAGE INPAINTING WITH SIMPLE USER GUIDANCE

Xin Zhang[1,2], *Bernd Hamann*[2], *Xiao Pan*[1], *Caiming Zhang*[1]

[1]School of Computer Science and Technology, Shandong University, Jinan, China
[2]Department of Computer Science, University of California, Davis, CA 95616, U.S.A.

ABSTRACT

We introduce a new approach for performing image inpainting, by devising an integrative method based on the superpixel segmentation technique and considering minimal user input. Image inpainting methods are concerned with filling in missing or replacing undesired regions in an image. Typically, inpainting methods consider and extrapolate known image data. Superpixels in the immediate neighborhood of the inpainting region are computed and used as source image data to fill in (or replace) the inpainting area. A user provides additional information by specifying line segments in the image to assist the otherwise automatic inpainting process, to ensure that only desirable superpixels are utilized when copying them into the inpainting region. User interaction is minimal, as it is merely necessary to specify a small number of line segments that define image parts to be used as source data in distinct inpainting regions. We provide experimental results demonstrating that our method performs well when compared against other methods, especially concerning the preservation of edges and texture in the inpainted regions.

Index Terms— image inpainting, superpixel, user guidance

1. INTRODUCTION

Image inpainting refers to the technique of completing missing or damaged areas in an image [1, 2], which is an important and active research topic in image processing. The development of image inpainting algorithms is driven by many applications, such as photo editing, architecture repair in images, 3D reconstruction and so on. As many contributions have been made to tackle image inpainting problem, we briefly review some of the main methods, i.e., the diffusion-based method, sparse representation based method and example-based method.

Diffusion-based method smoothly propagates local image structures from the exterior to the interior of the hole. This method is pioneered by Bertalmio et al. [1] with a successful scheme based on PDE (Partial Differential Equation) method. It is further extended by joint interpolation of vector fields and gray levels [3], curvature-driven diffusion model [4], total variational (TV) inpainting[5, 6], and so on. These methods are naturally well suited for completing small regions, such as straight lines and curves, but less effective in handling large missing regions due to the inability to synthesize textures.

The basic idea of image sparse representation is to represent image by sparse combination of an over-complete set of transformations. For inpainting problem, missing pixels are inferred by adaptively updating sparse representation [7, 8, 9]. This method may fail to recover image structures and introduce smooth effect when filling large missing regions.

Example-based methods are developed from texture synthesis [10, 11]. Missing textures are synthesized by sampling, copying and stitching together patches taken from the known part of the image. According to the number of candidate exemplars, it is further categorized into one-candidate [12, 13], multiple-candidates [14, 15] and global [16, 17] methods. The one-candidate method chooses the best match for each image patch. Multiple-candidate method infers the missing region using weighted average of multiple candidate patches. Global method defines inpainting as a global optimization problem. In [16], the objective function is optimized by priority belief propagation. Example-based inpainting methods have better performance for images with big holes, but generally suffer from high computational cost. Another limitation is the difficulty of synthesizing textures with perspective transformations and to fill in large and dispersed holes. Moreover, although performing better than diffusion methods on texture areas, the example-based methods often suffer from error propagation and repetitive patterns, which make images unnatural, especially in the case of stochastic textures.

To solve the problems in example-based method, we propose an efficient image inpainting method based on superpixel and simple user guidance in this paper. The contributions of our work are two-fold:

- We constrain source exemplars to neighbor superpixels to increase inpainting reliability. Searching efficiency for best exemplar is greatly improved.
- Simple user guidance is introduced to segment the source regions and lost regions correspondingly. Error

This work was supported by China Scholarship Council (201606220188), National Natural Science Foundation of China (61373078, 61572292, 61373088), NSFC Joint Fund with Zhejiang Integration of Informatization and Industrialization under Key Project (U1609218).

propagation is effectively avoided.

2. ALGORITHM

The example-based inpainting method proposed in [18] serves as one component of our framework. We first summarize some basic notions.

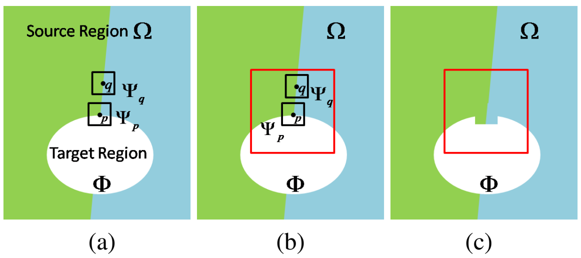


Fig. 1. Basic notions of image inpainting method.

As illustrated in Fig. 1(a), the region to be filled (white hole) is called target region Φ, and the known region is called source region Ω. The task of image inpainting is to complete the target region using the source region, and make the inpainted image satisfy a human's perception. The black-framed patch Ψ_p is the target patch with center pixel p. The similar patch in the source region is defined as Ψ_q with center pixel q. Missing pixels in patch Ψ_p are learned by sampling and copying the most similar patch Ψ_q from the source region, according to a certain distance metric. The key aspects of example-based inpainting method are inpainting order and distance metric for computing image patch similarity. We use the sparsity-driven order proposed in paper [19], which first fills edges and corners while moving from the boundaries inward. A patch Ψ_p with highest priority is inpainted first. The priority of patch Ψ_p is defined as

$$P(\Psi_p) = C(\Psi_p)Sparsity(\Psi_p), \tag{1}$$

where $C(\Psi_p)$ is patch confidence, i.e., the relative proportion of source pixels in the patch. The term $Sparsity(\Psi_p)$ is defined as

$$Sparsity(\Psi_p) = ||\vec{\omega}_p||_2\sqrt{N(\Psi_p)/N}, \tag{2}$$

where $\vec{\omega}_p$ is a vector of normalized similarities between Ψ_p and source patches in the neighbor window centered at pixel p, shown as red square in Fig. 1(b). $N(\Psi_p)$ is the number of source patches in the window, and N is the number of pixels in the window. The distance metric $d(\Psi_p, \Psi_{q_i})$ is defined as square root of the sum of squared difference (SSD) between patch Ψ_p and Ψ_{q_i}. The similarity between patch Ψ_p and Ψ_{q_i} is defined as

$$\omega_{p,q_i} = e^{-\frac{d(\Psi_p, \Psi_{q_i})}{\sigma^2}}. \tag{3}$$

The sparsity term defined in Eq. (2) measures whether a patch contains structures. If a patch contains structures, e.g., edges or corners, its similar patches are sparsely distributed along the structures. For plain patches, their similar patches are homogeneously distributed in the window. The L_2 norm of the similarity vector is higher for sparsely distributed similarity values [19]. Thus, the sparsity term of a patch containing structures is higher so as to make it inpainted first.

Compared with diffusion-based methods and sparse representation methods, the example-based method is more suitable for filling in relatively larger holes. However, as the source region is the whole known region, it is possible to produce flaws, such as those shown in Fig. 2(b), Fig. 4(c). Thus, we make further analysis and improve the method by introducing a new approach that combines the superpixel method with a simple user-guided technique to perform inpainting.

2.1. Superpixel-based image inpainting

For example-based methods, using the whole known region as exemplar source leads to error texture propagation, and is also quite time-consuming to search the best match patch. We constrain the best match patch by adding a location term L when defining the new distance metric as,

$$d\_new(\psi_p, \psi_{q_i}) = d(\psi_p, \psi_{q_i}) + \lambda L(\psi_p, \psi_{q_i}), \tag{4}$$

where $L(\psi_p, \psi_{q_i})$ is the Euclidean distance between two patch centers and $d(\psi_p, \psi_{q_i})$ measures the SSD of two patch intensities as in Eq. (3). λ is a weight factor, and we set it to one empirically. With the new distance metric, locally similar patches are preferred for inpainting missing pixels and the inpainting result is more satisfactory.

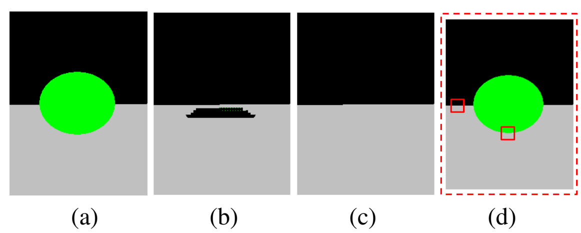
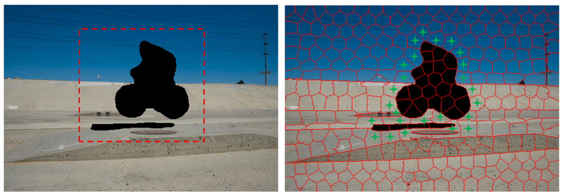


Fig. 2. (a) Image with green hole to be inpainted. (b) Example-based method [18]. (c) New distance metric. (d) Selection of similar patches of result (b).

As shown in Fig. 2, by using the new distance metric, the inpainting result in Fig. 2(c) is much more acceptable than the one in Fig. 2(b), where black pixels appear in the gray region. The reason is illustrated in Fig. 2(d). When matching patches, the known parts are compared. Thus, the two red-framed patches are perfectly matched. However, the black pixels in the source patch are not suitable for the missing part. Here, the upper part of the image should be black and the lower part should be gray. If the whole source region is used for searching, the patches across two colors will be chosen as best match because they are in front of the gray patches in the patch sequence. Thus, it is possible that false pixels are produced and propagated. With the new distance metric, a best match is constrained to the local area. In fact, it is intuitive that local information is more reliable for inpainting. Besides, it is computationally expensive to search the whole source region.



(a) (b)

Fig. 3. (a) Original image with black hole; (b) Superpixel segmented image. Superpixels marked with green stars are used for inpainting.

Based on the above analysis, we come to a conclusion that local information is more suitable for inpainting than the whole source region. Thus, we propose a superpixel-based image inpainting method. A superpixel is a cluster of image pixels with the same or similar properties, which is an over-segmenting technique for images. We adopt the method called FCCS [20] as a pre-processing step for image inpainting, which is efficient for good boundary adherence. As shown in Fig. 3(b), neighbor superpixels marked with green stars along the boundary of the black target region are used as best exemplar source. Compared with a regular window, such as the red dotted line in Fig. 3(a), superpixels are more flexible and adhere boundaries very well, which can hold wrong textures from propagating into the target region.

2.2. User guidance

Based on experience and highly developed perception skills, humans can understand complex images easily, and can intuitively determine what to do to inpaint a hole in a meaningful way. Humans can recognize different objects in the images and determine which target region belongs to which object, which in fact is the segmentation of images. Thus, when humans inpaint an image, we subconsciously segment the source image into different objects, and also segment the target region into corresponding regions. We use corresponding objects to fill different target regions. Inspired by this process, simple user guidance is introduced in our approach. Specifically, we let a user click an even number of points at the junction of different objects. Each pair of points determines a straight line to segment the target region and source region into two parts. Neighbor superpixels in the same segmentation are used to inpaint the corresponding target region.

Fig. 4 shows the effectiveness of user guidance. Fig. 4(b) is the result produced by the proposed method. The edges are clear, and no flaws appear. The result by the example-based method [18] in Fig. 4(c) is unpleasant. Even after modifying it with the location constraint according to Eq. (4), there are still obvious flaws in Fig. 4(d). Most flaws are produced by copying from improper regions. Thus, it is necessary to define boundaries to constrain the search region. With simple user guidance, this problem is easily solved, and the time required for specific points is negligible.

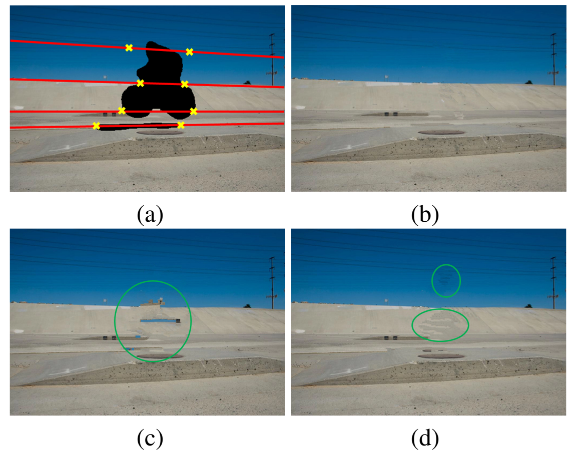
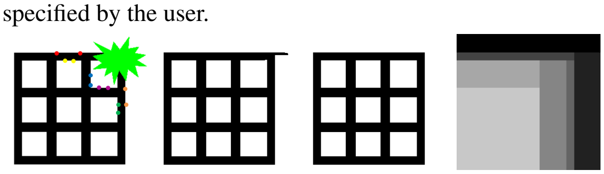


Fig. 4. Effectiveness of user guidance. (a)Eight points specified by user define four line segments; (b)Proposed method. (c)Example-based method [18]. (d)New distance metric.

Furthermore, with the help of user guidance, the corner can be perfectly recovered. Fig. 5(a) is a regular lattice with a corner covered by an irregular object. It is very difficult for computer to recover the corner without understanding the pattern. Fig. 5(b) shows the result by [18], and Fig. 5(c) is the result obtained with our new method, based on 12 points specified by the user.



(a) (b) (c) (d)

Fig. 5. Recovering corner for a regular pattern. (a) Original image with the upper-right corner lost. The same color points specified by user define one line segment. (b) Example-based method [18]. (c) Proposed method. (d) Segmentation of original image based on 12 points specified by user.

3. EXPERIMENT

In this section, we compare the proposed method with four methods, Bugeau et al. [12], Herling & Broll [21], Getreuer [22] and Xu & Sun [19]. Bugeau et al. combined PDE-based method with example-based method [12]. Method Herling &

Broll [21] inpaints images by minimizing a global cost function. Getreuer develped total variation inpainting method using split bregman [22]. The algorithm by Xu and Sun [19] is based on sparsity order and fills the target region by combining several patches. Test images and the mask for target region are from dataset published by paper [23]. The inpainting area accounts for 10% pixels. According to the image inpainting survey [23], the popular image quality metrics, such as PSNR and SSIM, are not suited to judge inpainting quality. Thus, we illustrate and analyze the experiment results from visual aspect.

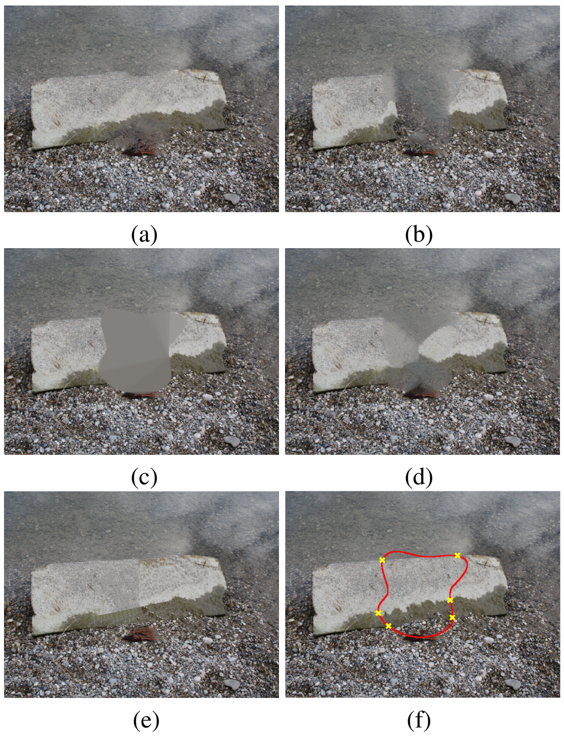


Fig. 6. (a)Bugeau et al.[12];(b)Herling&Broll[21];(c)Getreuer [22]; (d) XU&Sun [19]; (e)Proposed; (f)Ground Truth with red circled target region and yellow user input points.

As is shown in Fig. 6(a), method Bugeau et al. [12] has better results on keeping edges than Herling & Broll [21], Getreuer [22] and Xu & Sun [19]. Although the inpainted edges by method Bugeau et al. [12] are kept connected, they are not straight and sharp. The stones in Fig. 6(b) and Fig. 6(d) are all broken as textures around the stone are propagated into the stone area. The target regions are fuzzy without proper textures kept. The result by method Getreuer [22] in Fig. 6(c) keeps the whole target region incoherent with the image. The proposed method gives the clearest boundaries of the stone as shown in Fig. 6(e). In Fig. 7, the inpainting results obtained by each method are similar with Fig. 6. The edges of paving stone are blurred in Fig. 7(a), and broken in Fig. 7(b) and Fig. 7(d). Fig. 7(e) keeps the edges sharp and textures in different objects are in good order.

Compared with the above methods, the proposed method produces the most acceptable inpainting results. The experiment further confirms the effectiveness of our proposed method, as the superpixel segmentation and user guidance effectively reduce the error texture propagation and pattern repetition.

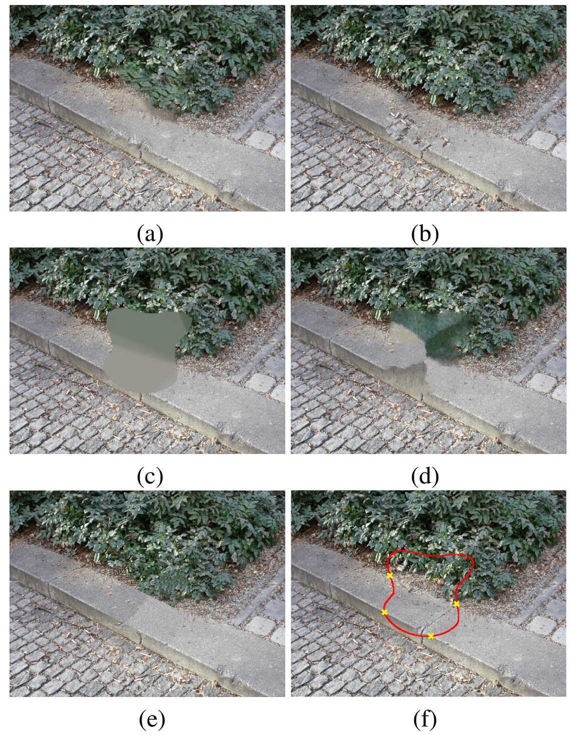


Fig. 7. (a)Bugeau et al.[12];(b)Herling&Broll[21];(c)Getreuer [22]; (d)XU&Sun[19]; (e)Proposed; (f)Ground Truth with red circled target region and yellow user input points.

4. CONCLUSION

In this paper, we propose an effective example-based image inpainting method. Rather than using pixels in the whole image, we adopt neighbor superpixels as sources, which makes the best matched patch more appropriate, and the algorithm efficiency are improved greatly. Furthermore, simple user guidance is introduced by specifying couples of points in the image to segment the source region and the target region. Only sources in the corresponding segmentation region are allowed to supply inpainting information. Cross-boundary flaws are effectively avoided, and the edges are well kept.

Our future work will focus on designing algorithm to detect brief line segmentation of the image, as automatic algorithm is more effective in practical application.

5. REFERENCES

[1] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester, "Image inpainting," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 417–424.

[2] Christine Guillemot and Olivier Le Meur, "Image inpainting: Overview and recent advances," *IEEE signal processing magazine*, vol. 31, no. 1, pp. 127–144, 2014.

[3] Coloma Ballester, Marcelo Bertalmio, Vincent Caselles, Guillermo Sapiro, and Joan Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE transactions on image processing*, vol. 10, no. 8, pp. 1200–1211, 2001.

[4] Tony F Chan and Jianhong Shen, "Nontexture inpainting by curvature-driven diffusions," *Journal of Visual Communication and Image Representation*, vol. 12, no. 4, pp. 436–449, 2001.

[5] T Chan and J Shen, "Local inpainting models and tv inpainting," *SIAM J. Appl. Math*, vol. 62, no. 3, pp. 1019–1043, 2001.

[6] Jianhong Shen and Tony F Chan, "Mathematical models for local nontexture inpaintings," *SIAM Journal on Applied Mathematics*, vol. 62, no. 3, pp. 1019–1043, 2002.

[7] Onur G Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*. IEEE, 2003, vol. 1, pp. I–713.

[8] Onur G Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part ii: adaptive algorithms," *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 555–571, 2006.

[9] Mohamed-Jalal Fadili, J-L Starck, and Fionn Murtagh, "Inpainting and zooming using sparse representations," *The Computer Journal*, vol. 52, no. 1, pp. 64–79, 2009.

[10] Alexei A Efros and Thomas K Leung, "Texture synthesis by non-parametric sampling," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*. IEEE, 1999, vol. 2, pp. 1033–1038.

[11] Alexei A Efros and William T Freeman, "Image quilting for texture synthesis and transfer," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM, 2001, pp. 341–346.

[12] Aurélie Bugeau, Marcelo Bertalmío, Vicent Caselles, and Guillermo Sapiro, "A comprehensive framework for image inpainting," *IEEE Transactions on Image Processing*, vol. 19, no. 10, pp. 2634–2645, 2010.

[13] Jino Lee, Dong-Kyu Lee, and Rae-Hong Park, "Robust exemplar-based inpainting algorithm using region segmentation," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, 2012.

[14] Olivier Le Meur, Josselin Gautier, and Christine Guillemot, "Examplar-based inpainting based on local geometry," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 3401–3404.

[15] Olivier Le Meur and Christine Guillemot, "Super-resolution-based inpainting," in *European Conference on Computer Vision*. Springer, 2012, pp. 554–567.

[16] Nikos Komodakis and Georgios Tziritas, "Image completion using efficient belief propagation via priority scheduling and dynamic pruning," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2649–2661, 2007.

[17] Yael Pritch, Eitam Kav-Venaki, and Shmuel Peleg, "Shift-map image editing," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 151–158.

[18] Antonio Criminisi, Patrick Pérez, and Kentaro Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on image processing*, vol. 13, no. 9, pp. 1200–1212, 2004.

[19] Zongben Xu and Jian Sun, "Image inpainting by patch propagation using patch sparsity," *IEEE transactions on image processing*, vol. 19, no. 5, pp. 1153–1165, 2010.

[20] Xiao Pan, Yuanfeng Zhou, Caiming Zhang, and Qian Liu, "Flooding based superpixels generation with color, compactness and smoothness constraints," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4432–4436.

[21] Jan Herling and Wolfgang Broll, "Pixmix: A real-time approach to high-quality diminished reality," in *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*. IEEE, 2012, pp. 141–150.

[22] Pascal Getreuer, "Total variation inpainting using split bregman," *Image Processing On Line*, vol. 2, pp. 147–157, 2012.

[23] Philipp Tiefenbacher, Viktor Bogischef, Daniel Merget, and Gerhard Rigoll, "Subjective and objective evaluation of image inpainting quality," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 447–451.