

Gradient Based Image Completion by Solving Poisson Equation

Jianbing Shen, Xiaogang Jin, Chuan Zhou

State Key Lab of CAD & CG, Zhejiang University, Hangzhou, 310027, P.R.China

Charlie C. L. Wang

*Department of Automation and Computer-Aided Engineering,
The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, P.R.China*

Abstract

This paper presents a novel gradient-based image completion algorithm for removing significant objects from natural images or photographs. Our method reconstructs the region of removal in two phases. Firstly, the gradient maps in the removed area are completed through a patch based filling algorithm. After that, the image is reconstructed from the gradient maps by solving a Poisson equation. A new patch-matching criterion is developed in our approach to govern the completed of gradient maps. Both the gradient and the color information are incorporated in this new criterion, so a better image completion result is obtained. Several examples and comparisons are given at the end of the paper to demonstrate the performance of our gradient-based image completion approach.

Key words: Image completion, Gradient, Filling order, Patch propagating, Poisson equation

1 Introduction

The removal of objects or the recovery of damaged portion in a given image, known as image completion, is an important task in the photo editing or the video post-processing. Given an input image I with a missing or unknown region Ω , the task of image completion is to propagate structure and texture information from the known region $I \setminus \Omega$ into Ω . The removed parts can be filled by various interactive tools such as clone brush strokes and compositing processes. However, automatically and seamlessly filling Ω by the known region in I is still a problem under research.

A number of approaches related to image completion have been proposed in computer graphics and computer vision literatures [2, 4–6]. In [4], Bertalmio *et al.* conducted a PDE-based method to repair damaged images. The idea is to extend the structures inwards by satisfying some continuities across the boundaries of the damaged area. The recovered results are highly smooth. This works well for small damaged regions. However, when the reconstructed area is large, it will give a blurry artifact that lacks texture. Therefore this approach fails in reconstructing large regions. Levin *et al.* [2] extended the idea by measuring global image statistics based on the prior image knowledge besides the local color information. Drori *et al.* [6] incorporated pyramid image approximation and adaptive image fragments to achieve impressive results. Nevertheless, all these approaches are extremely slow due to the high computational complexity.

Recently, some researchers have considered exemplar-based methods to complete images with large portion removed [3, 5, 6, 8, 9, 16, 19]. The first attempt to use exemplar-based synthesis for image completion was by Harrison [8], who filled the pixels in the target region by the level of "textureness" on the neighborhoods of a pixel. Although the intention sounds good, strong linear structures were often overruled by nearby noises in his approach. Jia and Tang [9] presented a technique for filling image regions by explicitly segmenting the unknown area into different homogeneous texture areas with a tensor voting method. However, their approach requires both an expensive segmentation step and a difficult choice about how to stitch the boundary between two textures. More recently, Criminisi *et al.* [5] proposed an exemplar-based image completion algorithm through region filling, where the patch filling order is determined by the angle between the isophote direction and the normal direction of the local filling front - so that the missing region with stronger structures can be filled in higher priority. Sun *et al.* [19] introduced a novel structure propagation approach to image completion. In their system, the user manually specifies important missing structures by extending a few curves or line segments from a known region to the unknown. Their approach synthesizes image patches along these user-specified curves using patches selected around the curves in the known region. Structure propagation is formulated as a global optimization problem by enforcing structure and consistency-constraints. If only a single curve is specified, structure propagation is solved using *Dynamic Programming*. When multiple intersecting curves are specified, the *Belief Propagation* algorithm is adopted to find the optimal patches. After completing the structure propagation, the remaining unknown regions are filled using patch-based texture synthesis.

None of these approaches [2–6, 8, 9, 16, 19] conducts gradient maps in the image completion. The approach presented in this paper recovers the unknown region Ω in two phases: in phase one, the region is filled in the gradient maps by a patch based best-first filling algorithm; in phase two, the image is recon-

structed in Ω from the reconstructed gradients by solving a Poisson equation.

The gradient-based image processing techniques and the Poisson equation solving techniques have been addressed in several related areas such as high dynamic range compression [7], Poisson image editing [11], image fusion for context enhancement [13], interactive photomontage [1], Poisson image matting [14] and photography artifacts removal [18]. In our approach, a new criterion is developed for selecting similar patches to fill the damaged region in gradient maps, so that the image can be reconstructed from gradients by solving a Poisson equation.

The contributions of our proposed approach consists of two aspects:

- A new framework for image completion is developed by integrating the techniques of 1) exemplar-based image completion, 2) image processing in gradient domain, and 3) Poisson equation;
- We introduce a better patch-matching criterion, which incorporates both the gradient and the color factors.

Based on this, as will be demonstrated later, our approach can generate better completion results than other existing methods in a comparable computing time.

2 Gradient-Based Image Completion

2.1 Algorithm Overview

The input of an image completion algorithm is an image I which contains a manually masked area as the unknown region Ω (the region to be filled), while the output is a modified image I' with Ω filled. Here, we conduct the similar notation as that used in the image completion literatures [4–6, 16, 19]. We denote the unknown region by Ω , the known region by Φ , the contour of Ω by $\partial\Omega$, the source and target patches by Ψ_s, Ψ_t respectively. The symbol $G(\dots)$ is conducted to represent the information and operations in the gradient domain. As mentioned above, the whole algorithm consists of two phases: 1) image inpainting of $G(\Omega)$ and 2) reconstruction of Ω from $G(\Omega)$.

The first phase is completed by iteratively copying the information from a source patch Ψ_s to a target patch Ψ_t , which shows the highest priority of patch-filling among all candidate patches with centers located on $\partial\Omega$ - the method for determining patch-filling order will be described in section 2.2. Ψ_s is a patch in Ψ showing the highest similarity to Ψ_t based on some matching

criterion (which will be detailed in section 2.3). After finding Ψ_s , the pixels $p \in \Psi_t$ with its confident parameter $C(p) = 0$ are filled by their corresponding pixels in Ψ_s . Note that the pixel values to be filled are its gradients $G(p)$ (both $G_x(p)$ and $G_y(p)$). Ω and $\partial\Omega$ should be updated correspondingly after filling, so as the priority of candidate patches on $\partial\Omega$. The purpose of the second phase is to compute a new I' based on the inpainted G so that $\|\nabla I' - G\|$ is minimized - this will be addressed in section 2.4.

In summary, the overall algorithm is listed below.

- (1) Initialization;
- (2) Selecting a target patch Ψ_t with its center $p_c \in \partial\Omega$, which is with the highest filling priority $P(p_c)$;
- (3) Searching Φ to determine a source patch Ψ_s , which shows the highest similarity between Ψ_s and Ψ_t ;
- (4) For $\forall p \in \Psi_t$, if its confident parameter $C(p) = 0$, filling its gradients $G(p)$ by the gradients $G(p')$ on its corresponding pixel $p' \in \Psi_s$ and setting $C(p) = 1$;
- (5) Updating Ω and $\partial\Omega$;
- (6) If $\Omega \neq \emptyset$, go back (2);
- (7) Computing I' which leads to $\min \|\nabla I' - G\|$.

During the initialization, the filling front $\partial\Omega$ is searched, and the confident parameter and the gradients on every pixel are computed.

2.2 Gradient-Confidence Based Patch-Filling Order

Image content features, such as color histogram, gradient, texture, shape and object composition, are usually introduced to extract image salient information. These features are usually employed to determine the filling order of patches into Ω . Zhang *et al.* [16] incorporated the texture in the neighborhoods of a patch to determine the filling order, and Criminisi *et al.* [5] used the angle between the isophote direction and the normal direction $\partial\Omega$ to define the priority of patch filling so that the patches along structures in the damaged region can be filled before filling other textures. In this paper we compute the patch filling order using image gradient feature which is akin to [5].

The confidence parameter $C(p)$ at a pixel p is initialized to zero if $p \in \Omega$; otherwise it is assigned to one.

$$C(p) = \begin{cases} 0, & \forall p \in \Omega \\ 1, & \forall p \in I \setminus \Omega \end{cases} \quad (1)$$

The second relevant parameter defined on p is the gradient term $G(p)$, which reflects the local shape feature, and its value is based on the magnitude of the gradient information at location p . The gradient term $G(p)$ is defined as follows:

$$G(p) = \frac{1}{|A|} \sum_{q \in A} \sqrt{G_x^2(q) + G_y^2(q)} \quad (2)$$

where A denotes the neighborhood area around pixel p , and $G = [G_x, G_y]$ denotes the gradient field of an image for the horizontal and vertical directions using a simple forward difference. Here, A is chosen with the same size as the patches Ψ_s and Ψ_t .

For a given patch Ψ_t centered at the point p_c for some $p_c \in \partial\Omega$, we define its filling priority as the following formula:

$$P(p_c) = \Upsilon(p_c) \cdot G(p_c), \quad (3)$$

$$\Upsilon(p_c) = \frac{\sum_{q \in \Psi_t \cap (I \setminus \Omega)} C(q)}{|\Psi_t|}, \quad (4)$$

where $|\Psi_t|$ represents the area of Ψ_t . The gradient term $G(p_c)$ is a measurement of the amount of edge and structure information surrounding the pixel p_c . The purpose for calculating the patch priority value $P(p_c)$ is to encourage the structural regions, which are with greater $P(p_c)$, to be filled first. Therefore this can help to propagate the broken lines into the connected ones. The greater value returned by $P(p_c)$, the higher filling priority is given on the patch Ψ_t .

2.3 Patch-Matching Criterion

After computing the patch filling priorities on $\partial\Omega$, the patch Ψ_t with the highest priority is firstly selected to fill gradients. The target region is filled with corresponding gradients on some source region Ψ_s . Therefore, the criterion of selecting a best source patch takes a central role in an exemplar-based image completion algorithm. As aforementioned, under the assumption that the content in the unknown area is similar to the content of the known region for some similarity measurement [16], the traditional inpainting techniques propagate pixel-information via diffusion (e.g., [4, 6]), which results in blurry artifact and discontinuous lines in large regions. Criminisi *et al.* [5] propagated the filling patches by directly sampling the source patch. Alternatively, here we only fill the gradients sampled from a source patch. In addition, we do not employ the *Sum of Squared Difference* (SSD) in [5], which is widely used in image completion to measure the similarity between space patches. The reason is that the SSD does not always suffice to provide the desired completion results (ref. [15]).

We develop an exponential similarity measurement to search the target patch Ψ_t in Φ , which incorporates both the color differences and the gradient differences between two patches. The similarity is defined as

$$s(\Psi_s, \Psi_t) = e^{d_c(\Psi_s, \Psi_t) + d_g(\Psi_s, \Psi_t)} \quad (5)$$

with

$$d_c(\Psi_s, \Psi_t) = \sum_{(x,y)} \|\Psi_s^c(x, y) - \Psi_t^c(x, y)\| \quad (6)$$

$$d_g(\Psi_s, \Psi_t) = \sum_{(x,y)} \|\Psi_s^g(x, y) - \Psi_t^g(x, y)\| \quad (7)$$

where Ψ_s^c, Ψ_t^c represent the color information of the source patch and the target one in RGB, while Ψ_s^g, Ψ_t^g represent the corresponding gradient vectors. For those unknown pixels, zeros are used as pixel values. Based on this new criterion, the source patch Ψ_s for filling the gradients in Ψ_t can be determined by

$$\Psi_s = \arg \min_{\Psi_i \in \Phi} \frac{s(\Psi_i, \Psi_t)}{|\Psi_i|} \quad (8)$$

The functionality of this new patch-matching criterion will be demonstrated later.

As described in the overall algorithm, after iteratively filling the gradients in Ω , we obtain the final updated gradient map $G' = [G'_x, G'_y]$ which is conducted to reconstruct the image I' by solving a Poisson equation.

2.4 Image Reconstruction from Gradients

Image reconstruction from gradient fields is a very active research area. In 2D, a modified gradient vector field $G' = [G'_x, G'_y]$ may not be integrable. Let I' denote the completion image reconstructed from G' , we employ one of the direct methods recently proposed in [7] to minimize $\|\nabla I' - G'\|$, so that $G \approx \nabla I'$. By introducing a Laplacian and a divergence operator, I' can be obtained by solving the Poisson differential equation [12, 18]

$$\nabla^2 I' = \text{div}([G'_x, G'_y]) \quad (9)$$

Since both the Laplacian ∇^2 and div are linear operators, approximating them using standard finite differences yields a large system of linear equations. We use the *full multigrid* method [10] to solve the Laplacian equation with Gaussian-Seidel smoothing iterations. This leads to $O(n)$ operations to reach an approximate solution, where n is the number of pixels in the region Ω . Note that the Poisson equation is only solved in Ω .

For solving the Poisson equation more efficiently, an alternative is to use a *rapid Poisson solver*, which uses a sine transform based on the method [12] to

invert the Laplacian operator. However, the complexity with the *rapid Poisson solver* will be $O(n(\log(n)))$. Therefore, the full multigrid method is employed in our implementation.

The regions to be filled are zero-padded on all sides, so that the Dirichlet boundary condition instead of the Neumann boundary condition is used to reconstruct the image I' , so that the scale-shift ambiguity mentioned in [18] is avoided.

3 Experimental Results and Discussions

Our algorithm has been applied to a variety of colorful photographs with complex background structures. As the purpose of image completion is to fill damaged areas while satisfying visual perception, it is commonly accepted that the quality of results is detected by the human perception of the appearance in completed images [6]. The experimental results demonstrate that satisfactory completed images have been generated by our approach. All the examples shown in this section are tested on a PC with Pentium IV 1.6GHz CPU + 512MB RAM. Similar to other patch-filling based image completion approaches (e.g., [5]), we employ patches with 9×9 pixels in all our examples except the one in Figure 7, where a better result is given on the patches with 13×13 pixels as larger patches make it more robust to the ambiguity on depth.

Our first example shown in Figure 1 is downloaded from the web site — <http://www.cis.rit.edu/fairchild>. The result shows that even if about 39% region (the foreground persons) has been removed, our method can still recover the background reasonably good in the damaged region. However, the result from [6] blurs the texture and lacks detail texture in the filled area. Moreover, our algorithm takes only about 4 minutes to compute the result while Drori *et al.*'s algorithm [6] requires more than 4 hours by our implementation.

Figures 2 and 3 show the comparisons of results obtained by our gradient-based method vs. other previous methods [4–6], where our method performs better. For the example II in Figure 2, it is easy to find the blurry artifact introduced by diffusion and the lack of texture in the result of [4] while ours shows no blurry effect. In example III (see Figure 3), a larger object has been removed, where our approach dramatically outperforms earlier works in terms of perceptual quality. The image generated by our approach provides more details and is more coherent than the results from other approaches. Notice that the comparison is based on our implementations of previous techniques [4–6]. One interesting observation is that the result of our implementation on [5] (see Figure 3(e)) is not as good as the one shown in their paper (ref. Figure 15(e) in [5]). If our new patch-matching criterion is applied in the

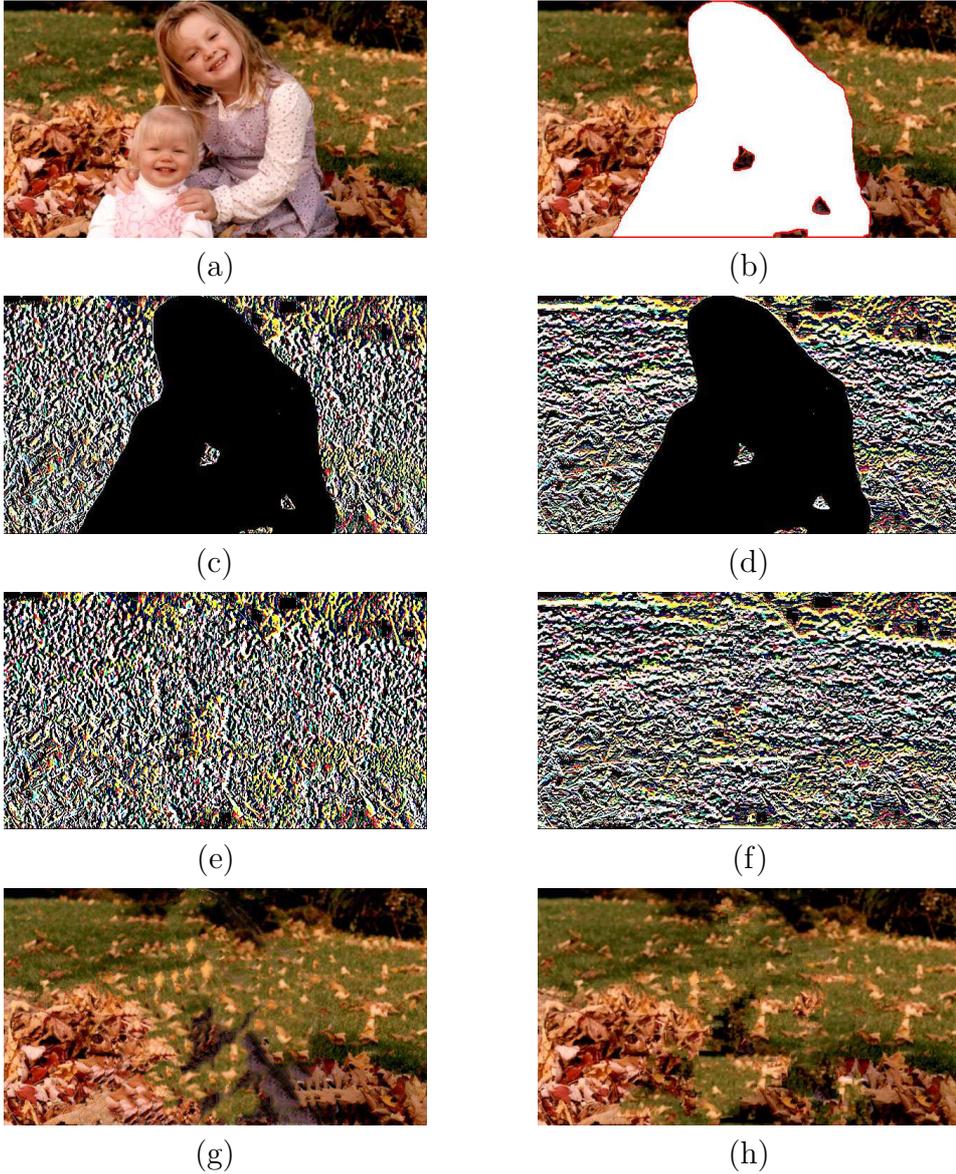


Fig. 1. Example I: (a) original image; (b) the figure needs to be completed (in white with red boundary); (c), (d) the initial gradient maps in horizontal and vertical directions; (e), (f) the inpainted gradient maps in horizontal and vertical directions; (g) the result of [6]; (h) the result by our method.

implementation of [5], the result is somewhat improved (see Figure 4(a)). A similar comparison is given on the performances of SSD and our criterion in our own image-completion framework (see Figure 4(b) vs. 4(c)), which also shows the advantage of our patch-matching criterion. The second comparison that could be given is about the color-based approach vs. gradient-based approach (see Figure 4(a) vs. 4(c)). It is easy to find that the gradient-based approach outperforms the color-based one.

Figure 5 shows the results of our gradient-based algorithm on one example



Fig. 2. Example II: (a) original image; (b) the region of removed microphone needs to be completed; (c) the result obtained by the inpainting method [4]; (d) the result determined by our method; (e) zoom-view of the marked region in the result of [4]; (f) zoom-view of our result.

used in [6]. It is obvious that, comparing the result from [6], our algorithm does not introduce the edge blur which is given on the result of [6]. In Figure 6, we complete the same image but in different filling orders - one is with the naive line-scan filling order while another is with the filling order suggested by the priority value (Eq.(3)). As has already been proved in [5], the scheme with order suggested by priority values gives a better result.

Considering about the computational time aspect, the cost of our method is much faster than [6], faster than [4], and comparative with [5]. Note that in the implementation of [4], we iterate 100,000 times. Ours is a little bit slower than [5] since we need to take some time to solve the Poisson equation. In our current implementation, the Poisson equation can be quickly solved in several seconds. The computing time statistics have been listed in Table 1.

The gradient-based method works well if the damaged structures can be

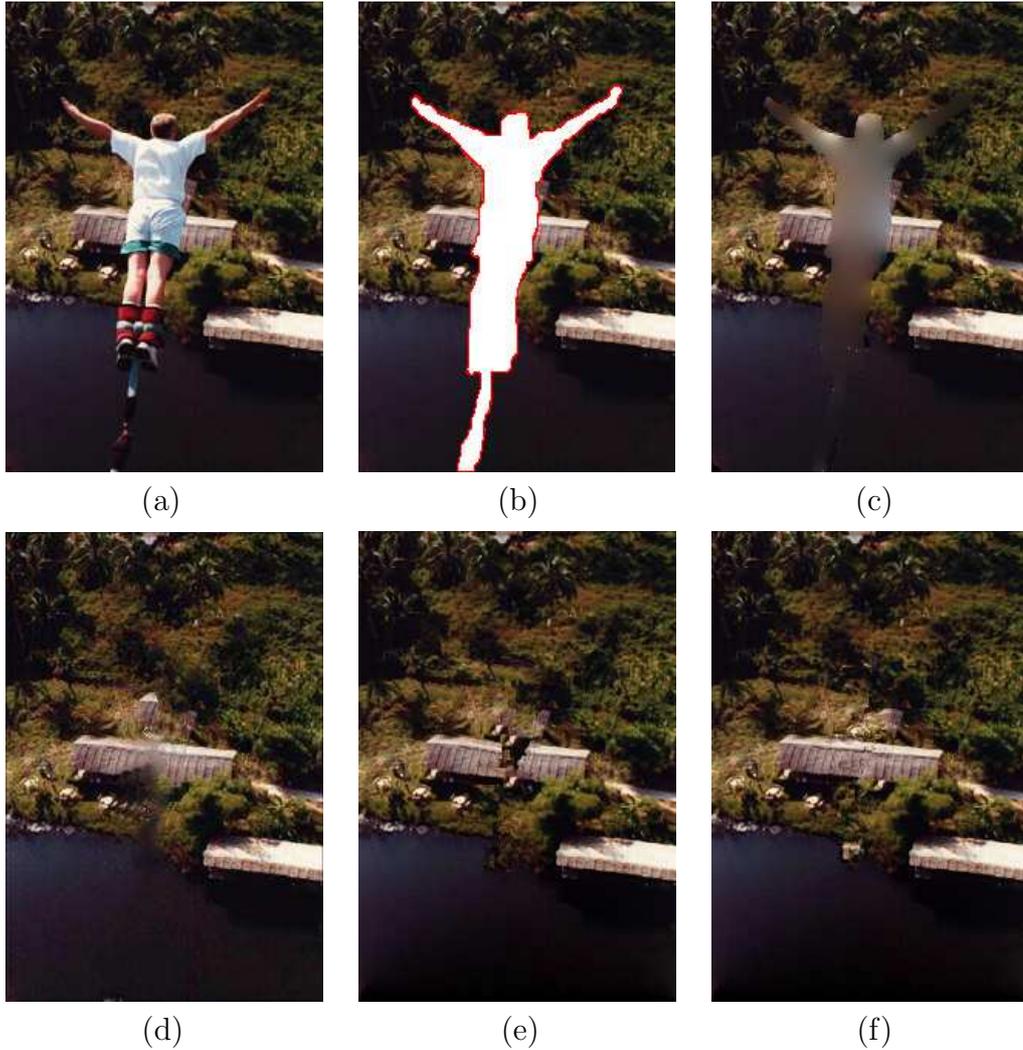


Fig. 3. Example III: (a) original image from [4]; (b) the removed region covers 13% of the total image area; (c) the result of region filling by traditional image inpainting [4]; (d) result from [6]; (e) the result image by exemplar-based completion [5] based on our implementation; (f) the final image where the bungee jumper has been completely removed and the occluded region has been well reconstructed by our new algorithm.

represented by a set of linear structures remained. However, our approach also shares the common limitations with other exemplar-based approaches: 1) if there are not enough samples in the given image, it will be impossible to reconstruct the desired structure or texture; 2) our approach has no ability to handle depth ambiguities, where the missing area covers the intersection of two perpendicular regions as shown in Figure 7; 3) our algorithm also has problems when dealing with curved structures. The later two limitations are partially solved in [19] by interactive user inputs. In our algorithm, pixel colors are represented in RGB, we may achieve better experimental results in the CIE Lab color space because of its property of perceptual uniformity (ref. [5, 21]),

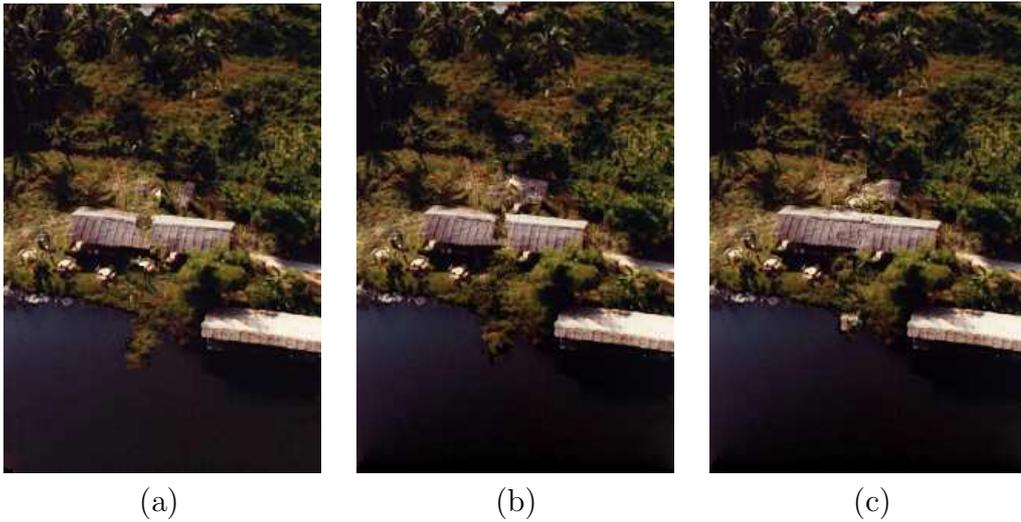


Fig. 4. Illustration for showing the performance of our patch-matching criterion and the gradient-based approach: (a) the completion by [5] but using our new patch-matching criterion; (b) the result from our gradient-based algorithm but with the SSD criterion; (c) the result of our full approach.

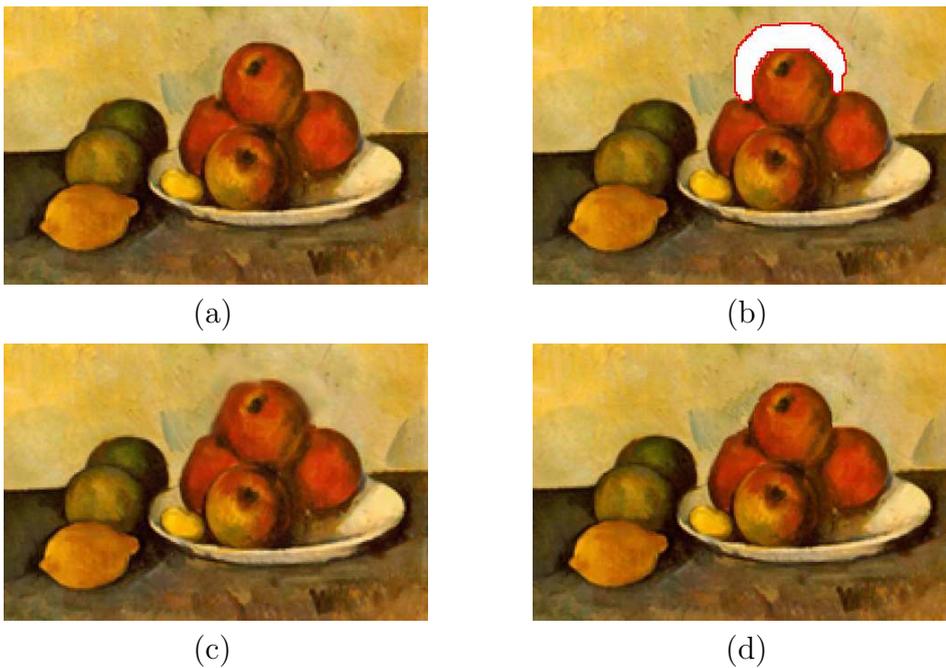


Fig. 5. Example IV: (a) a photo of the oil painting “Still Life with Apples”, by P. Cézanne, c. 1890; (b) the manually selected region to be completed; (c) result obtained by [6]; (d) result obtained by our gradient method. Note that our result is more sharp.

which leads the similarity distances more meaningful than in RGB.



Fig. 6. The results from the naive line-scan patch filling order (left) vs. the priority based patch filling order (right).

Table 1
Computational Statistics

Examples	Fig. 1	Fig. 2	Fig. 3
Image Size	600×338	480×640	206×308
Inpainting Area	78,824 pixels	3,235 pixels	7,997 pixels
Time of [4]	261.78 min.	7.82 min.	18.24 min.
Time of [5]	4.23 min.	1.02 min.	0.81 min.
Time of [6]	4.17 hr.	2.38 hr.	2.67 hr.
Time of ours*	4.47 min.(1.03 sec.)	1.01 min.(1.97 sec.)	1.08 min.(1.28 sec.)

*The time listed in bracket is the time for solving Poisson equations.

4 Conclusions and Future Work

A novel gradient-based image completion algorithm by solving Poisson equation has been proposed in this paper. Our image completion approach is conducted in two phases. At first, the gradient maps of the damaged region are reconstructed by an exemplar-based method which depends on a new patch-matching criterion incorporating both the color and the gradient factors. After determining the gradient maps, the completion result is computed from the gradients by solving a Poisson equation. Experimental results demonstrate both the feasibility and the efficiency of our algorithm.

Currently, we are intending to extend our approach from the still photography completion to the areas of video completion and the completion of mesh surfaces (e.g., [20]). The difficulties in removing objects from video include the global motion compensation and the method to maintain consistency of the unknown area over the whole video sequence [15, 17, 22].

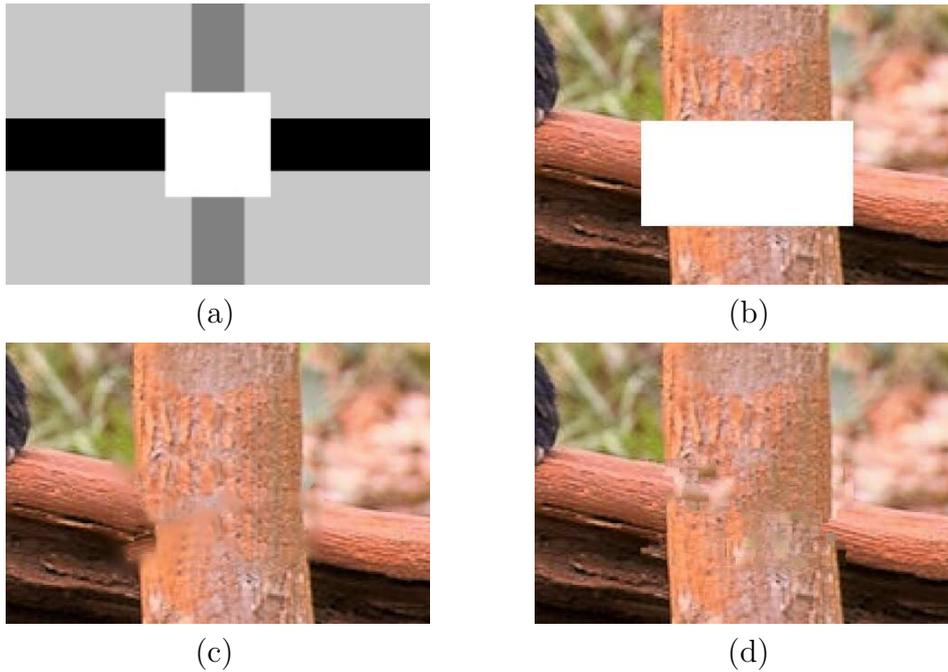


Fig. 7. Our approach does not handle depth ambiguities: (a) a synthesis image in which the missing area covers the intersection of two perpendicular regions [6]; (b) a nature image with the same ambiguity; (c) the result obtained by [6]; (d) the result obtained by our method.

5 Acknowledgements

This work was supported by 973 Program (No.2002CB312101), National Natural Science Foundation of China (No. 60273054, 60340440422) and Fok Ying Tung Education Foundation (No. 91069).

References

- [1] Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S., Colburn, A., Curless, B., Salesin, D., and Cohen, M. Interactive digital photomontage. *ACM Transactions on Graphics* 2004;23(3):294-302.
- [2] Levin, A., Zomet, A., and Weiss, Y. Learning how to inpaint from global image statistics. In: *Proceedings of IEEE International Conference on Computer Vision '03*, France. 2003. p. 305-361.
- [3] Bornard, R., Lecan, E., Laborelli, L., and Chenot, J. H. Missing data correction in still images and image sequences. In: *Proceedings of ACM Multimedia '02*, France, 2002. p. 355-361.

- [4] Bertalmio, M., Sapiro, G., Ballester, C., and Caselles, V. Image inpainting. In: Proceedings of SIGGRAPH '00, New Orleans, 2000. p. 417-424.
- [5] Criminisi, A., Pérez, P., and Toyama, K. Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing* 2004;13(9):1200-1212.
- [6] Drori, I., Cohen-Or, D., and Yeshurun, H. Fragment-based image completion. *ACM Transactions on Graphics* 2003;22(3):303-312.
- [7] Fattal, R., Lischinski, D., and Werman, M. Gradient domain high dynamic range compression. *ACM Transactions on Graphics* 2002;21(3):249-256.
- [8] Harrison, P. A non-hierarchical procedure for re-synthesis of complex texture. In: Proceedings of International Conference in Central Europe Computer Graphics and Visualization '01, Czech Republic, 2001. p. 190-197.
- [9] Jia, J., and Tang, C. K. Image repairing: Robust image synthesis by adaptive tensor voting. In: Proceedings of Conference on Computer Vision and Pattern Recognition '03, Madison, WI, 2003. p. 643-650.
- [10] Press, W., Teukolsky, S., and Vetterling, W., and Flannery, B. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press; 1992.
- [11] Pérez, P., Gangnet, M., and Blake, A. Poisson image editing. *ACM Transactions on Graphics* 2003;22(3):313-318.
- [12] Raskar, R., Tan, K., Feris, R., Yu, J., and Turk, M. Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging. *ACM Transactions on Graphics* 2004;23(3):679-688.
- [13] Raskar, R., Ilie, A., and Yu, J. Image fusion for context enhancement and video surrealism. In: Proceedings of Non-Photorealistic Animation and Rendering '04, France, 2004. p. 85-95.
- [14] Sun, J., Jia, J., Tang, C.K., and Shum, H.Y. Poisson matting. *ACM Transactions on Graphics* 2004;23(3):315-321.
- [15] Wexler, Y., Shechtman, E., and Irani, M. Space-time video completion. In: Proceedings of Conference on Computer Vision and Pattern Recognition'04, Washington, D.C., USA, 2004. p. 120-127.
- [16] Zhang, Y.J., Xiao, J.J., and Shah, M. Region completion in a single image. EUROGRAPHICS, Grenoble, France, Short Presentations. 2004.
- [17] Zhang, Y.J., Xiao, J.J., and Shah, M. Motion layer based object removal in videos. In: Proceedings of IEEE Workshops on Application of Computer Vision '05, Breckenridge, Colorado, USA, 2005. p. 516-521.
- [18] Agrawal, A., Raskar, R., Nayar, S.K., and Li, Y. Removing flash artifacts using gradient analysis. *ACM Transactions on Graphics* 2005;24(3):828-835.

- [19] Sun, J., Yuan, L., Jia J., and Shum, H.Y. Image completion with structure propagation. *ACM Transactions on Graphics* 2005;24(3):861-868.
- [20] Sharf, A., Alexa, M., and Cohen-Or, D. Context-based surface completion. *ACM Transactions on Graphics* 2004;23(3):878-887.
- [21] Kasson, J.M., and Plouffe, W. An analysis of selected computer interchange color spaces. *ACM Transactions on Graphics* 1992;11(4):373-405.
- [22] Jia, Y.T., Hu, S.M., and Martin, R.M. Video completion using tracking and fragment merging. *The Visual Computer* 2005;21(8-10):601-610.