

Iterative Gradient-Driven Patch-Based Inpainting

Sarawut Tae-o-sot¹ and Akinori Nishihara²

¹ The Department of Communications and Integrated Systems,
Tokyo Institute of Technology, Tokyo, Japan
`sarawutt@nh.cradle.titech.ac.jp`

² The Center for Research and Development of Educational Technology,
Tokyo Institute of Technology, Tokyo, Japan
`aki@cradle.titech.ac.jp`

Abstract. A novel exemplar-based image inpainting is proposed in this paper. This method is based on iterative approach which provides better result than greedy one. The problem of inconsistent results caused by raster scanning on target patch selection in iterative approach is focused in this paper. The proposed gradient-driven ordering is used to select target patch instead of traditionally predefined ordering. Due to the information-driven nature, this new approach is image's rotation invariant which means the same result is provided by different rotation of the same damaged image. Moreover, a random search approach is redesigned to be more reasonable and suitable for our novel gradient-driven ordering. The proposed method provides the best inpainting result among several well-known exemplar-based inpainting techniques including both greedy and iterative approach.

Keywords: image completion, image inpainting, exemplar-based, patchmatch.

1 Introduction

Image inpainting is the research area in the field of image processing whose goal is to remove some objects or restore damaged regions in a way that observers cannot notice the flaw. There are many applications of image inpainting such as photo editing, video editing, image compression and image transmission. Generally image inpainting techniques can be categorized into two approaches, Diffusion-based and Exemplar-based approach.

Diffusion-based approach is the fundamental approach in which information diffuses from known region into missing region. The problem is usually modelled by Partial Differential Equation (PDE), so sometimes it is called a PDE-based approach. Bertalmio *et al.* [2] reconstructed missing regions by diffusing known region along isophote direction, the direction of equal intensity value, into the missing region by a heat flow model. Chan *et al.* [4] introduced Total Variance (TV) framework for inpainting problem, then curvature-driven diffusion (CDD)

[5] which fixes connectivity problem in TV model. Diffusion-based approach works well for non-texture image, in which the missing region must be small and thinner than the surrounding object. In the case that the missing region is large or containing texture, this approach gives a blurry result.

Exemplar-based approach is originated from the Exemplar-based texture synthesis in [7] which synthesizes texture by copying the best match patch from the known region. However, directly applying exemplar-based texture synthesis to image inpainting problem may not provide satisfactory result. This is because, there are both structures and textures in natural images. Bertalmio [3] proposed to decompose the image into structural and textural images, then apply diffusion-based inpainting to the structural image and texture synthesis to the textural image separately. The result of combining restored structural and textural image is better than restoration by only diffusion-based inpainting or texture synthesis alone. However, that technique still cannot recover the large missing region. Criminisi *et al.* [6] introduced patch priority, which is defined by isophote direction and known pixels in target patch, for exemplar-based texture synthesis to determine the fill-in order. In that way, the structural information is recovered because the target patches which have high structural information are likely to be filled first. Kwok *et al.* [10] introduced DCT-based inpainting in which patch matching process is done in DCT domain. In that way, the error which is caused by noise is reduced by the noise reduction properties of DCT. However, new error is produced by the gradient-based filling process which roughly approximates the unknown region of the target patch before doing DCT. Patch shifting approach [13] was introduced for enhancing [6] and [10] performance. The main idea of the approach is to simultaneously reduce the unknown region and increase the known region of each target patch.

The major drawback of greedy approaches mentioned earlier is that filled region cannot be refilled. The error of the prior filled patches affect the next filled patch. This means that the error is accumulated. Due to this difficulty, iterative approach was considered. Komodakis *et al.* [9] modelled image inpainting problem by Markov Random Field and introduced priority belief propagation for solving. Graph cuts technique was introduced to optimize the problem in the work of Lee *et al.* [11]. Exemplar-based inpainting proposed by Wexler *et al.* [14] modelled inpainting as global optimization problem. Unlike [6], unknown region is filled iteratively until the solution converges. Generally, the iterative approach provides better results by sacrificing the computational time. However, patchmatch technique [1], which is a fast and efficient tool for dealing with exemplar-based problems, can reduce this huge complexity. A fast computational time is the result of random search strategy which compromises with the final result. For the best result of patchmatch inpainting, the structure of damaged area need to be manually specified. Unlike most of the greedy approaches, target patches are in the raster scan order which leads to inconsistency results. This means restored results of the same damaged image with different rotation may be different.

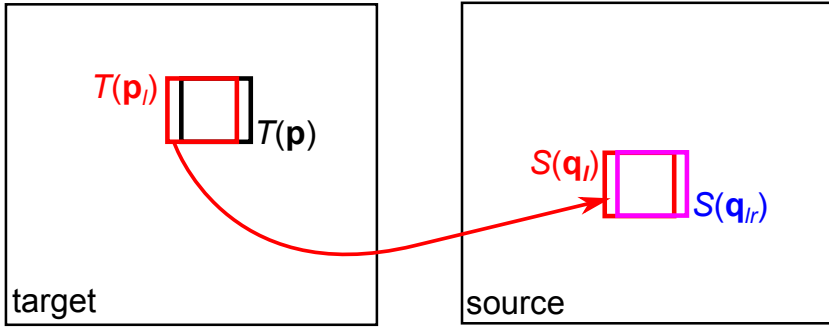


Fig. 1. Propagation step: candidate suggested from left neighbour of $T(\mathbf{p})$

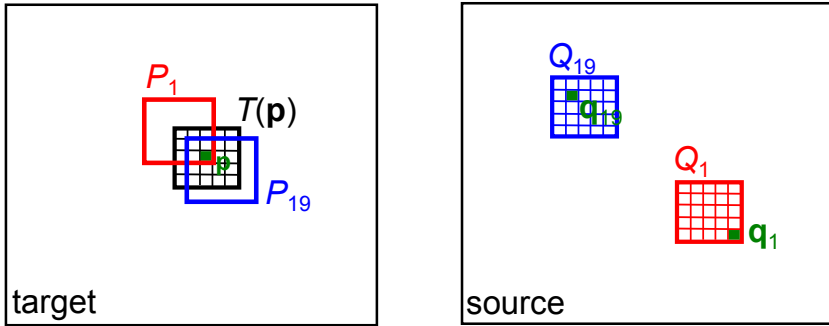


Fig. 2. Image reconstruction: Example of how corresponding pixel is defined

In this paper, iterative gradient-driven exemplar-based image inpainting, which is inspired by a fast but inconsistent patchmatch method, is proposed. Gradient-driven ordering is introduced in our proposed method. With this ordering, fill-in orders are arranged by the strength of gradient in target patch instead of traditional raster scanning. In other word, the patch with high structural information receives a priority to do patch matching first. Moreover, random search strategy is redesigned to be more efficient and able to support gradient-driven ordering scheme. By our novel framework the better and more consistent result can be achieved with almost the same speed as patchmatch inpainting.

This paper is organized as follows. In section 2, a new random search strategy is presented. In section 3, Gradient-driven ordering is introduced. Image reconstruction from nearest-neighbour field is explained in section 4. The experimental result of our framework and several inpainting techniques are presented and discussed in section 5. Finally, the conclusion of our work and the direction of our work are mentioned in section 6.

2 Random Search Strategy

Generally image inpainting is modeled as the problem of filling-in the missing region Ω , sometimes called target region, of the given image domain U by the information of the known region, sometimes called source region $U \setminus \Omega$. In exemplar-based view point, this is the problem of searching for the best match patch of all target patch T , where $T \in \Omega$, from the source region $U \setminus \Omega$. Exhaustive search was used in [14] and the results were excellent. Undoubtedly, computational complexity of this method is very high which is mathematically $O(mM_\Omega M)$, where m , M_Ω and M are numbers of pixels in patch, unknown region, Ω , and known region, Φ , respectively. The faster search method was introduced in [1]. The complexity of this method is around $O(mM_\Omega \log M)$ and can be optimized to $O(\sqrt{m}M_\Omega \log M)$. Cooperating with multi-resolution image processing patchmatch can be implemented in almost real-time.

Our search strategy is inspired by [1], however, many details are changed in order to make it compatible with our gradient-driven ordering scheme. Basically, the search strategy is the combination of 3 steps including initialization, propagation and random search.

2.1 Initialization

In this stage, all the patches in unknown region are assigned its corresponding patch randomly. In case of multi-resolution processing, the corresponding patches are initialized by the upscaled version of the previous low resolution level.

2.2 Propagation

Target patch $T(\mathbf{p})$ will consider the 4 candidate patches from the suggestion of its 4-neighbour patches including left neighbour $T(\mathbf{p}_l)$, right neighbour $T(\mathbf{p}_r)$, upper neighbour $T(\mathbf{p}_u)$ and lower neighbour $T(\mathbf{p}_d)$. The candidate of each neighbour is suggested from a viewpoint of its corresponding patch. For example, considering the left neighbour $T(\mathbf{p}_l)$ in Fig. 1, if $S(\mathbf{q}_l)$ is the corresponding patch of $T(\mathbf{p}_l)$, it would be possible that the right neighbour of $S(\mathbf{q}_l)$ is the corresponding patch of $T(\mathbf{p})$. So, $S(\mathbf{q}_{lr})$, the blue patch, is chosen as the candidate. In the same manner, $T(\mathbf{p}_r)$, $T(\mathbf{p}_u)$ and $T(\mathbf{p}_d)$ will suggest $S(\mathbf{q}_{rl})$, $S(\mathbf{q}_{ud})$ and $S(\mathbf{q}_{du})$ as candidates. Finally, the new corresponding patch of $T(\mathbf{p})$ can be calculated from

$$T(\mathbf{p}) = \arg \min_S \{d(T(\mathbf{p}), S(\mathbf{q}_{lr})), d(T(\mathbf{p}), S(\mathbf{q}_{rl})), d(T(\mathbf{p}), S(\mathbf{q}_{ud})), d(T(\mathbf{p}), S(\mathbf{q}_{du})), d(T(\mathbf{p}), S(\mathbf{q}))\}, \quad (1)$$

where $d(T, S)$ is Euclidean distance of patch T and S .

2.3 Random Search

This step randomly picks up the candidate from the searching region whose radius exponentially decreases. Unlike [1], the center of searching region may



Fig. 3. The experiment on restoring known image: (a) the original image, (b) the damaged image, (c) the result of Criminisi's method with PSNR = 35.42 dB, (d) the result of DCT method with PSNR = 36.17 dB, (e) the result of Wexler's method with PSNR = 37.28 dB, (f) the worst result of patchmatch inpainting with PSNR = 38.21 dB, (g) the best result of patchmatch inpainting with PSNR = 39.15 dB, (h) the result of our proposed method with PSNR = 39.35 dB

not be fixed to the center of the corresponding patch from propagation step \mathbf{q}_0 . At every decrease in searching radius a new corresponding patch is determined by the previous corresponding patch $S(\mathbf{q}_{i-1})$ and the random candidate $S(\mathbf{q}_i)$. If the corresponding patch changes, the center for randomly selecting the candidate patch will also change. This procedure would be repeated until the search radius is less than 1. A complete strategy of this step can be described as follows,

initialization the center of searching region is set to the center of corresponding patch \mathbf{q}_0 .

iteration

-new candidate is chosen by

$$\mathbf{q}_i = \mathbf{q}_{i-1} + w\alpha^i \mathbf{R}_i, \quad (2)$$

where w is the maximum search radius, α is reduction ratio of search radius and \mathbf{R}_i is a uniform random vector in $[-1, 1] \times [-1, 1]$. Normally, α is set to 0.5.

-the new corresponding patch is updated by

$$S(\mathbf{q}_i) = \arg \min_S \{d(T(\mathbf{p}), S(\mathbf{q}_{i-1})), d(T(\mathbf{p}), S(\mathbf{q}_i))\} \quad (3)$$

With this novel strategy, the corresponding patch seems to converge faster than the traditional strategy. A random candidate patch in each iteration always locates around the corresponding patch of each iteration. In contrast, the random candidates of [1] are densely around candidate patch of propagation step $S(\mathbf{q}_0)$. This condition may impede the convergence.

3 Gradient-Driven Ordering

On exemplar-based image inpainting, a predefined filling order, such as raster scan or onion peel, usually leads to unsatisfactory result as discussed in [6]. Moreover, this also leads to an inconsistent result. Not only greedy approach but also iterative approach, which is affected by predefined filling order. As shown in Fig. 3(f) and (g) and 4, with different raster scan directions, different results are achieved. In other word, the reconstruction results of same damaged image but different rotations, e.g., the damaged image and its 90° rotation, are probably different. And we do not know which rotation would provide the best result.

In this paper, structural information of patch is involved in ordering process. Magnitude of gradient is used as a measure of structure. Generally, on human point of view, a good restored image should contain no discontinuous structure [8]. To maintain this criterion, the patch with high magnitude of gradient should have a high priority to be selected as a target patch. Anyway, reliability of target patch must also be considered. So, the patches on the boundary of unknown region, which should be more reliable than ones inside the unknown region, must be considered first.

The concept of Gradient-driven ordering is to select the patch on the boundary of unknown region with the highest magnitude of gradient first as target patch. Then searching for its corresponding patch by random search strategy in section 2. After the target patch get its correspondence, the new boundary is defined by ignoring the target patch. And new target patch is selected again from this new boundary. This procedure is repeated until all patches in unknown region have

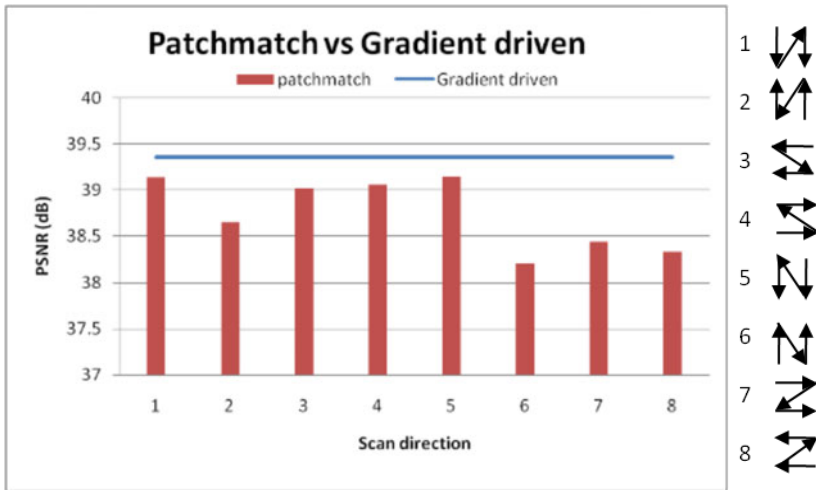


Fig. 4. PSNR of restoration of 3(b) by all scanning direction of patchmatch vs. our proposed method

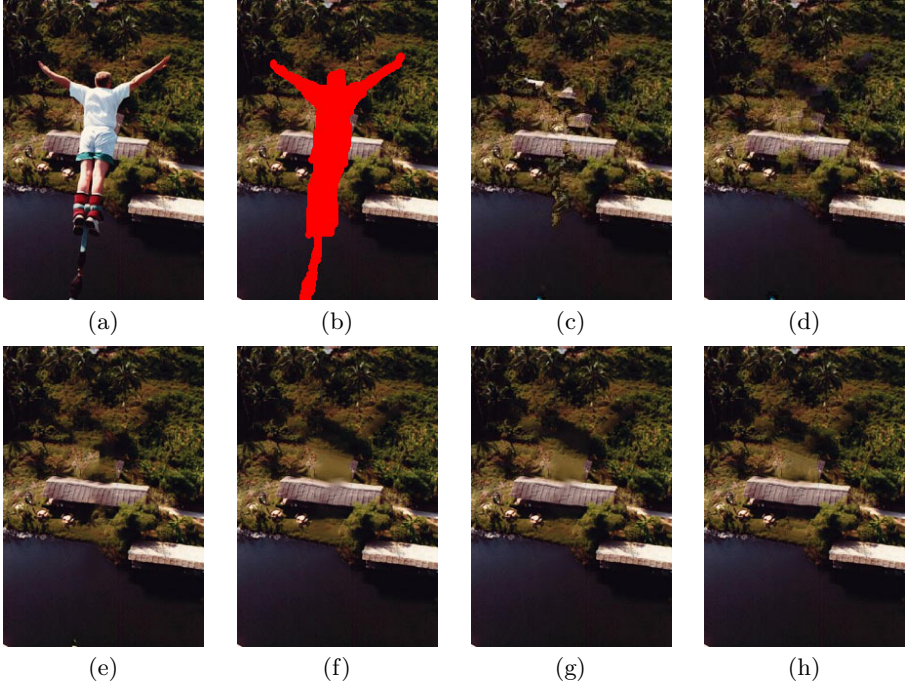


Fig. 5. The experiment on removing undesired object: (a) the original image of bungee jumping man, (b) removed region, (c) the result of Criminisi’s method, (d) the result of DCT method, (e) the result of Wexler’s method, (f) result of patchmatch inpainting with vertical raster scan from top-left to bottom-right corner, (g) result of patchmatch inpainting with horizontal raster scan from top-right to bottom-left corner, (h) the result of our proposed method

their corresponding patch. The whole process of gradient-driven ordering can be described as follows,

- initialization
 - Set all known pixels as “freeze”.
 - Set all unknown pixels as “active”.
- iteration
 - Step 1 Calculate magnitude of gradient on the boundary of “active” region.
 - Step 2 Select patch with maximum magnitude of gradient as target patch $T(\mathbf{p}_{max})$.
 - Step 3 Search for the corresponding patch of $T(\mathbf{p}_{max})$.
 - Step 4 Set \mathbf{p}_{max} as “freeze”.
 - Step 5 If there is still “active” pixel left, go to step 1.

4 Image Reconstruction

After finishing random search procedure for all $T(\mathbf{p}_i)$ where $\forall \mathbf{p}_i \in \Omega$, each $T(\mathbf{p}_i)$ then has its corresponding patch $S(\mathbf{q}_i)$. This section shows how to reconstruct the unknown region from all corresponding patches $S(\mathbf{q}_i)$. Denote P_1, P_2, \dots, P_M as all the patches that contain \mathbf{p} of $T(\mathbf{p})$. The number of patch M is equal to the size of the patch. Q_1, Q_2, \dots, Q_M denote the corresponding patches of P_1, P_2, \dots, P_M . And $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_M$ denote the corresponding pixel of \mathbf{p} suggested by Q_1, Q_2, \dots, Q_M respectively. For example, the patch Q_1 would suggest \mathbf{q}_1 for \mathbf{p} and Q_{19} would suggest \mathbf{q}_{19} for \mathbf{p} as shown in Fig. 2. In general, the best solution should compromise all the suggestion. So, the restored intensity $\hat{I}(\mathbf{p})$ should minimize the error from all suggestions. This error $e(I(\mathbf{p}))$ can be defined as follows,

$$e(I(\mathbf{p})) = \frac{1}{M} \sum_{i=1}^M (I(\mathbf{q}_i) - I(\mathbf{p}))^2 \quad (4)$$

After differentiating $e(I(\mathbf{p}))$ with respect to $I(\mathbf{p})$ and equating to zero, $I(\mathbf{p})$ can be obtained from

$$\hat{I}(\mathbf{p}) = \frac{1}{M} \sum_{i=1}^M I(\mathbf{q}_i). \quad (5)$$

This is the update equation for pixel \mathbf{p} . To reconstruct the unknown region all pixel inside the region need to be updated by this equation. Actually, $\hat{I}(\mathbf{p})$ is an average of all the intensity of the candidate pixel $I(\mathbf{q}_i)$. Equation (5) is just the simple version of update equation in [12] where completeness term is ignored.

5 Experimental Results and Discussion

In our experiment, beside showing the restored image, numerical evaluation is also used. In the case where the original image is available, peak signal to noise ratio (PSNR) is used to evaluate the result. Multi-resolution processing is applied to all the implementation of our proposed method.

5.1 Uniqueness of the Proposed Method and Inconsistency of Iterative Approach

In this experiment, the effect of raster scanning direction is investigated. All raster scanning directions have been implemented. In Fig. 3(f) and (g), the result of patchmatch inpainting with 7×7 patch size but different scanning direction are shown. Fig. 3(f), whose PSNR is 38.21 dB, is the worst result of all scanning direction. This image is achieved by scanning on vertical direction from bottom-left to top-right of the image. Fig. 3(g), whose PSNR is 39.15 dB, is the best result of all scanning direction. This image is achieved by scanning on vertical direction from top-right to bottom-left of the image. Obviously, these

two direction are just the opposite of each other. And for patchmatch algorithm, these are the same process with switching scan order. These results show that the result of patchmatch is very sensitive to scanning direction. With gradient-driven ordering on 7×7 patch size, the result, shown in Fig. 3(*h*), looks better than 3(*f*) and (*g*). And PSNR of Fig. 3(*h*), which is 39.35dB, is also higher than 3(*f*) and (*g*). According to Fig. 4, the proposed method outperforms patchmatch inpainting on all directions.

Another example is shown in Fig. 5(*f*)-(*h*). The result of our proposed method, shown in Fig. 5(*h*), obviously look better than the results of patchmatch inpainting, shown in Fig. (*f*) and (*g*). And, also, the results of patchmatch inpainting are not unique when the scanning direction is changed.

5.2 The Performance of the Proposed Method vs. Other Well-Known Inpainting Methods

In this experiment, the performance of our proposed method is compared with several well-known inpainting techniques except patchmatch inpainting which has already been discussed in the previous experiment. Firstly, the artificial damaged image shown Fig. 3(*b*), generated from the original image Fig. 3(*a*), is used for evaluating the performances. Because of the availability of original image, the performances can be evaluated in both visual and numerical aspect. The best result of Criminisi's method with 9×9 patch size is shown in Fig. 3(*c*). PSNR of the result is 35.42dB. Fig. 3(*d*) shows the best result of DCT method where PSNR is 36.17 dB. Although PSNR of DCT method is higher than Criminisi's method, it is obviously seen that the result of Criminisi's method has more natural look. So, just only numerical result may not measure the performance of inpainting. The result of Wexler's method is shown in Fig. 3(*e*) whose PSNR is 37.28. In this image, some discontinuity can be noticed as seen on the shoulder of the model. The result of our proposed method with 7×7 patch size is shown in Fig.3(*h*). Our proposed method achieved the highest PSNR of 39.35 dB. Visually, Fig. 3(*h*) seems the best reconstructed image. The shoulder of the model is restored perfectly by our method while other techniques fail.

In Fig. 5, the performance for object removing task is shown. The Bungee jumping man, which is masked as red region in Fig. 5(*b*), on the original image in Fig. 5(*a*) would be removed by various techniques. Fig. 5(*c*)-(*e*) and (*f*) show the best result by Criminisi's method, DCT method, Wexler's method and our proposed method respectively. Some discontinuity with strong false edge can be noticed on the result of Criminisi's method with 5×5 patch size in Fig. 5(*c*). The result of smooth edge can be noticed in DCT method with 9×9 patch size as shown in Fig. 5(*d*). In Fig. 5(*e*), Wexler's method with 9×9 patch size gives the result with uneven intensity and a little discontinuity of edge. The result of our proposed method with 7×7 patch size has the best reconstructed structure (a roof top) as shown in Fig. 5(*h*). However, there is noticeable blur at texture region (on the top of the roof). This effect may be the result of average updating equation (5) and fixed patch size. In other word, 7×7 patch size may be suitable for reconstructing structure of Fig. 5(*b*) but the smaller patch size may be needed

for reconstructing texture region. This assumption will be investigating on our future works.

6 Conclusion

This paper proposes a novel iterative version of exemplar-based image inpainting. A main contribution of our proposed method is that the target patch is selected by consideration of structural information. Unlike traditional raster scan approach, whose results are inconsistent, the structures of the damage image tend to be reconstructed perfectly and consistently by our approach. This means that the reconstruction results of the same damaged image with different rotation are unique by our method. Moreover, A fast computation speed of our inpainting method is achieved by an integration of a redesigned random search strategy. Our proposed method outperforms several well-known exemplar-based inpainting on both greedy and iterative approaches, especially on structural region. However, in some cases, the texture nearby the reconstructed structure is flattened. As our assumption, the patch which suitable for reconstruct structure may not be suitable for reconstruct texture. So, designing for inpainting algorithm with adaptive patch size is our future work.

References

1. Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D.B.: PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics (Proc. SIGGRAPH)* 28(3) (August 2009)
2. Bertalmío, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: *SIGGRAPH*, pp. 417–424 (2000)
3. Bertalmío, M., Vese, L.A., Sapiro, G., Osher, S.: Simultaneous structure and texture image inpainting. *IEEE Transactions on Image Processing* 12(8), 882–889 (2003)
4. Chan, T.F., Shen, J.: Local inpainting models and tv inpainting. *SIAM Journal on Applied Mathematics* 62(3), 1019–1043 (2001)
5. Chan, T.F., Shen, J.: Non-texture inpainting by curvature-driven diffusions (cdd). *J. Visual Comm. Image Rep.* 12, 436–449 (2001)
6. Criminisi, A., Pérez, P., Toyama, K.: Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing* 13(9), 1200–1212 (2004)
7. Efros, A., Leung, T.: Texture synthesis by non-parametric sampling. In: *International Conference on Computer Vision*, pp. 1033–1038 (1999)
8. Kanizsa, G.: *Organization in Vision*. Holt, Rinehart Winston (1979)
9. Komodakis, N., Tziritas, G.: Image completion using efficient belief propagation via priority scheduling and dynamic pruning. *IEEE Transactions on Image Processing* 16(11), 2649–2661 (2007)
10. Kwok, T.H., Sheung, H., Wang, C.C.: Fast query for exemplar-based image completion. *IEEE Transactions on Image Processing* 19, 3106–3115 (2010)
11. Lee, S.Y., Heu, J.H., Kim, C.S., Lee, S.U.: Object removal and inpainting in multi-view video sequences. *International Journal of Innovative Computing, Information and Control* 6(3(B)) (March 2010)

12. Simakov, D., Caspi, Y., Shechtman, E., Irani, M.: Summarizing visual data using bidirectional similarity. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2008, pp. 1–8 (June 2008)
13. Tae-o-sot, S., Nishihara, A.: Exemplar-based image inpainting with patch shifting scheme. In: 17th International Conference on Digital Signal Processing (2011)
14. Wexler, Y., Shechtman, E., Irani, M.: Space-time completion of video. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 463–476 (2007)