DCT Inpainting with Patch Shifting Scheme

Sarawut TAE-O-SOT* and Akinori NISHIHARA†
* The Department of Communications and Integrated Systems, Tokyo Institute of Technology, Tokyo, Japan
E-mail: sarawutt@nh.cradle.titech.ac.jp
† The Center for Research and Development of Educational Technology, Tokyo Institute of Technology, Tokyo, Japan
E-mail: aki@cradle.titech.ac.jp

***Abstract*—This paper presents a novel exemplar-based image inpainting whose patch matching process is done in frequency domain in order to gain the advantage of noise reduction property. Moreover, a simple patch shifting scheme is also introduced to make sure that the target patch contains enough information to refer the unknown region. Each target patch is iteratively shifted to the position where there are enough known pixels on the patch as defined by user. This scheme can improve the traditional exemplar-based inpainting techniques which do not consider on the number of known and unknown pixels in the target patch. This ignorance can ruin the inpainting result when few known pixels are in the target patch. More meaningful target patches are provided by the proposed scheme. Experimental results show a significant improvement from traditional exemplar-based approach in both visual and mathematical aspects. Sharper and more continuous edge can be achieved by our approach.**

I. 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 approaches.

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.* [1] 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.* [2] introduced Total Variation (TV) framework for inpainting problem, then curvature-driven equation (CDD) [3] 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 of Efros and Leung [4]. In their work, the texture is synthesized 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 [5] proposed to decompose the image into structural and textural images, then applied 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 the known region in the 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.* [7] 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. More sophisticated exemplar-based inpainting was proposed by Wexler *et al.* [8]. It modelled inpainting as global optimization problem. Unlike [6], unknown region is filled iteratively until the solution converges. A fast iterative exemplar-based inpainting method which is called Patchmatch [9] was proposed by Barnes. 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.

Comparing with diffusion-based inpainting, exemplar-based approach gives a better result even in the large missing region case. However, in some cases, satisfactory results cannot be achieved because large unknown region is filled by small number of known pixels. That situation usually occurs although the number of known pixels is one criterion for choosing the target patch (as it involves in the confident term). In this paper, we apply patch shifting technique, which would provide more informative and reliable target patch, to DCT inpainting. This technique gave quite satisfactory result on our previous work [10]. And it promises to work more efficiently with DCT inpainting because the effect of rough estimation of unknown region on target patch is reduced.

This paper is organized as follows. In section II, the idea of exemplar based image inpainting is discussed. In section III, our proposed technique is presented. The experimental result of our technique and the traditional inpainting are presented

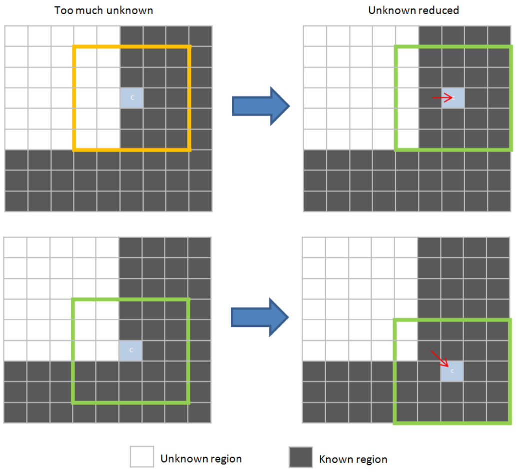

Fig. 1. The idea of patch shifting.

and discussed in section IV. Finally, we conclude our work in section V.

II. Exemplar-based Approach

A. General idea of exemplar-based inpainting

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 inpainting, in order to fill the target patch $\Psi_{\mathbf{p}}$, which is centered at pixel $\mathbf{p}$ and partially within Ω, the best match patch $\Psi_{\hat{\mathbf{q}}}$, which is centered at $\hat{q}$, is chosen from the source region. Then the intensities of $\Psi_{\mathbf{p}}$ in the target region are filled by copying from the corresponding pixels of $\Psi_{\hat{\mathbf{q}}}$. The order of selecting target patch intensively affects the restored result as the example shown in [6]. For a natural looking result, the edges should be continued which means the patch which contains high structural information should be filled first. With this principle, patch priority P is introduced. It is determined by the magnitude of the isophote direction and the known pixel in the patch. On each iteration, the target patch which has the highest patch priority is filled. Mathematically, patch priority is defined as

$$P(\mathbf{p}) = C(\mathbf{p})D(\mathbf{p}). \tag{1}$$

The confident term $C(\mathbf{p})$ and data term $D(\mathbf{p})$ are defined as follows;

$$C(\mathbf{p}) = \frac{\sum_{\mathbf{q}\in\Psi_{\mathbf{p}}\bigcap U\setminus\Omega} C(\mathbf{q})}{\mid \Psi_{\mathbf{p}} \mid} \quad \text{and} \quad D(\mathbf{p}) = \frac{\mid \nabla I_{\mathbf{p}}^{\perp} \cdot \mathbf{n}_{\mathbf{p}} \mid}{\alpha}, \tag{2}$$

where $\mid \Psi_{\mathbf{p}} \mid$ is the number of pixels of patch $\Psi_{\mathbf{p}}$, $\mathbf{n}_{\mathbf{p}}$ is the normal vector of the front $\partial\Omega$, $\nabla I_{\mathbf{p}}^{\perp}$ is the isophote at $\mathbf{p}$ and α is the normalizing factor which equals 255 for 8-bit grey-scale image. In our implementation, $\mathbf{n}_{\mathbf{p}}$ is unit vector of gradient of mask image M where $M(\mathbf{p}) = 1$ for $\forall\mathbf{p} \in \Omega$ and $M(\mathbf{p}) = 0$ for $\mathbf{p} \in U \setminus \Omega$. And $\nabla I_{\mathbf{p}}^{\perp}$ is computed from the maximum image gradient in $\Psi_{\mathbf{p}} \cap I$. The confident term C shows the ratio of known pixels which surround the center of the target patch. The data term D shows the strength of the edge at the target patch.

The process of Exemplar-based approach can be described as follows. Firstly, the confident term is initialized by assigning to $C(\mathbf{p}) = 0$ for $\forall\mathbf{p} \in \Omega$ and $C(\mathbf{p}) = 1$ for $\mathbf{p} \in U \setminus \Omega$. Then the following processes are repeated until the filling front $\partial\Omega^t = \emptyset$.

1. Identify the filling front $\partial\Omega$.
2. Compute patch priorities of all the patches whose center align on filling front $\partial\Omega$.
3. Chose the patch $\Psi_{\mathbf{p}}$ which has the maximum patch priority.
4. Find the best match patch $\Psi_{\hat{\mathbf{q}}}$ of $\Psi_{\mathbf{p}}$ from the source region $U \setminus \Omega$.
5. Copy data from $\Psi_{\hat{\mathbf{q}}}$ to $\Psi_{\mathbf{p}}$ for $\forall\mathbf{p} \in \Psi_{\mathbf{p}} \bigcap \Omega$.
6. Set $A = \Psi_{\mathbf{p}} \bigcap \Omega$, then update $\Omega = \Omega \setminus \Psi_{\mathbf{p}}$.
7. Update $C(\mathbf{p})$ for $\forall\mathbf{p} \in A$.

Note that, the best match patch $\Psi_{\hat{\mathbf{q}}}$ in step 4 is the patch which minimizes the Sum of Squared Differences (SSD) between itself and $\Psi_{\mathbf{p}}$ in known region. SSD is defined as

$$d(\Psi_{\mathbf{p}}, \Psi_{\mathbf{q}}) = \sum_{(i,j)} \mid I(\mathbf{p}_{(i,j)}) - I(\mathbf{q}_{(i,j)}) \mid^2, \quad (\forall\mathbf{p}_{(i,j)} \in \Psi_{\mathbf{p}} \text{ and } \forall\mathbf{q}_{(i,j)} \in \Psi_{\mathbf{q}}). \tag{3}$$

B. Extension to frequency domain

The exemplar-based approach can be applied to both conventional inpainting (spatial domain) [6] and DCT-based inapainting [7]. In the case of DCT-based inpainting, $I(\mathbf{p}_{(i,j)})$ and $I(\mathbf{q}_{(i,j)})$ in (3) are the DCT of the target patch and the candidate patch respectively. High frequency component of the patch can be ignored while computing SSD in order to reduce the influence of noise.

Target patch need to be filled in order to do DCT. In this paper, we propose to fill the unknown region of the target patch by traditional structural inpainting [1].

III. Patch Shifting

As we discussed previously that the best result of exemplar-based approach may not be achieved because in some cases the target patch has not enough known pixels for a meaningful representation. This situation can occur and ruin the final result, although the number of known pixels is a parameter to consider on the patch priority. In Fig. 1, it obviously seems that the target patches on the right column would produce better result than the target patches on the left column. In this paper, we introduce an easy but efficient approach to modify the target patch in the way that it always contains enough known pixels to produce more reliable result. Our idea is to shift the target patch to the known region in the case that there are not enough known pixels in that patch.

As shown in the first row of Fig. 1, if 15 known pixels (60% of the patch size) is not enough for the criteria then the

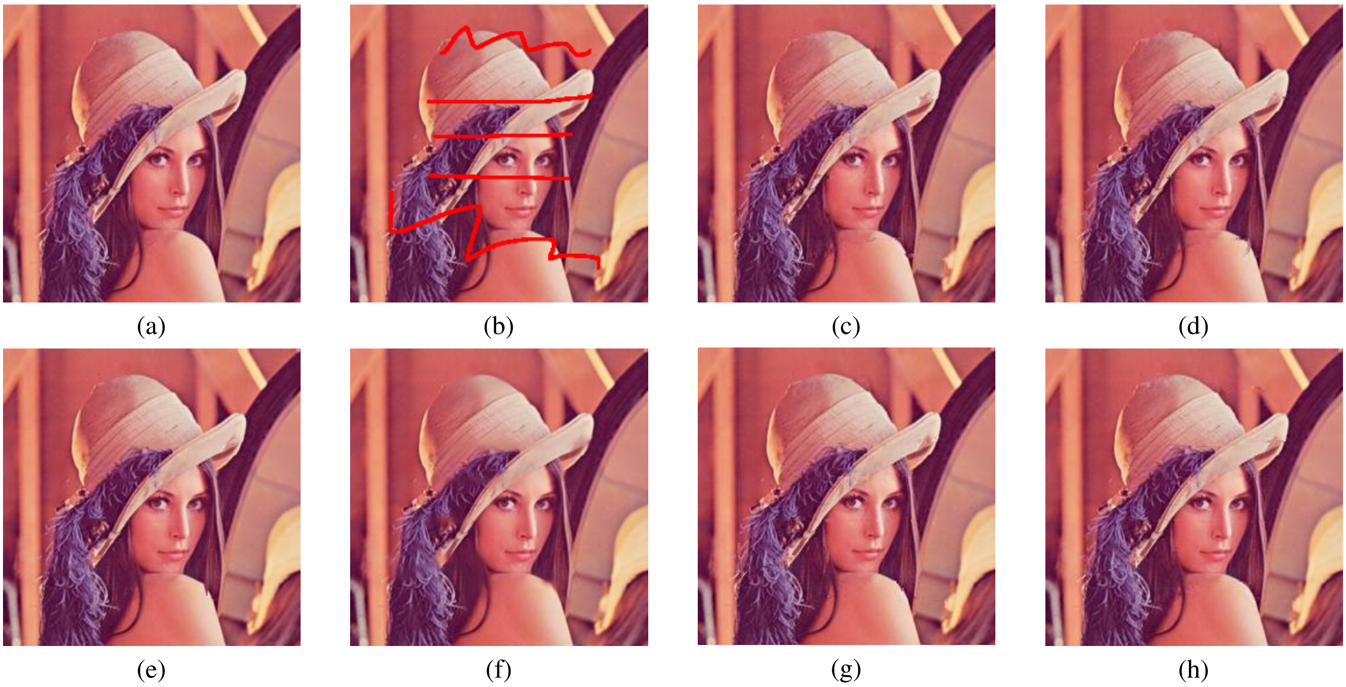

Fig. 2. The experiment on removing of known image: (a) the original image, (b) the damaged image, (c)the result of Criminisi's method with PSNR = 35.42 dB, the result of DCT method with PSNR = 36.17 dB, (e) the result of Wexler's method with PSNR = 37.28 dB, (f) the result of patchmatch inpainting with PSNR = 35.83 dB, (g) the result of patch shifting which applying to Criminisi's method, PSNR = 36.24 dB, (h) the result of proposed method with PSNR = 37.12 dB.

target patch is shifted to the right as shown on the right. In this case, we gain 5 more known pixels (20% of the patch size). For more clear understanding, let us consider the second row of Fig. 1. If known pixels must be more than 76% of patch size in target patch, the patch should be shifted to the lower right direction as shown in bottom right of Fig. 1. After shifting, we gain 4 more pixels (16% of patch size). In practice, we can find the minimum vertical shift S_v and horizontal shift S_h of the patch by

$$S_v(\mathbf{p}) = \sum_{n=-1}^{1} \sum_{n=-1}^{1} \psi(i+m, j+n) V_{m+2,n+2},$$
$$S_h(\mathbf{p}) = \sum_{n=-1}^{1} \sum_{m=-1}^{1} \psi(i+m, j+n) H_{m+2,n+2}, \quad (4)$$

where

$$\mathbf{V} = \begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} +1 & 0 & -1 \\ +1 & 0 & -1 \\ +1 & 0 & -1 \end{bmatrix}, \quad (5)$$

ψ is a mask image whose pixel is 0 at known pixel and 1 at unknown pixel, and $\mathbf{p} = (i, j)$ is the center of the target patch. $\mathbf{V}$ and $\mathbf{H}$ are called vertical and horizontal shifting penalty matrix respectively. They are designed under assumption that target patch should be shifted to the opposite direction of the unknown pixel. The 8-neighbour of the target patch would vote for the opposite direction against it if it is an unknown pixel. For example, considering on $\mathbf{V}$, it gives penalty $+1$ to the 3-upper neighbours of target center $\mathbf{p}(i, j)$ which is unknown pixel. And penalty -1 is given to the 3-lower neighbours of $\mathbf{p}(i, j)$ which is unknown pixel. Finally the voting direction or the the minimum vertical shift S_v would be given by the sum of all penalty on the 8-neighbour of $\mathbf{p}(i, j)$. The minimum horizontal shift S_h can also be obtained in the same way as S_v.

On exemplar-based inpainting, patch shifting is applied to the target patch with maximum priority whose number of known pixels is less than the prescribed threshold. The target patch repeatedly shifts by shifting vector $[S_v, S_h]$ obtained from (3) until the number of known pixel is more than the threshold. Then, the best matched patch of the shifted target patch is searched. However, if the promised target patch cannot be achieved while none of the pixel in shifted patch is in the initial patch, the next target patch with lower maximum priority is chosen and do patch shifting again. These processes are done repeatedly until satisfied target patch is found. Note that, to maintain the advantages of patch priority, we apply patch shifting to only limited number of target patches. In our experiment, patch shifting is applied to the first 100 maximum priority patch. If there is no satisfied patch, the shifted patch of the maximum priority patch is chosen.

IV. Experimental Results and Discussion

In this section, we do some experiments to evaluate the performance of our proposed method. Firstly, we investigate on how damaged image is recovered. The damaged image of Fig. 2(a) shown in Fig. 2(b) is inpainted by various techniques. The best result of Criminisi's method with 9×9 patch size is

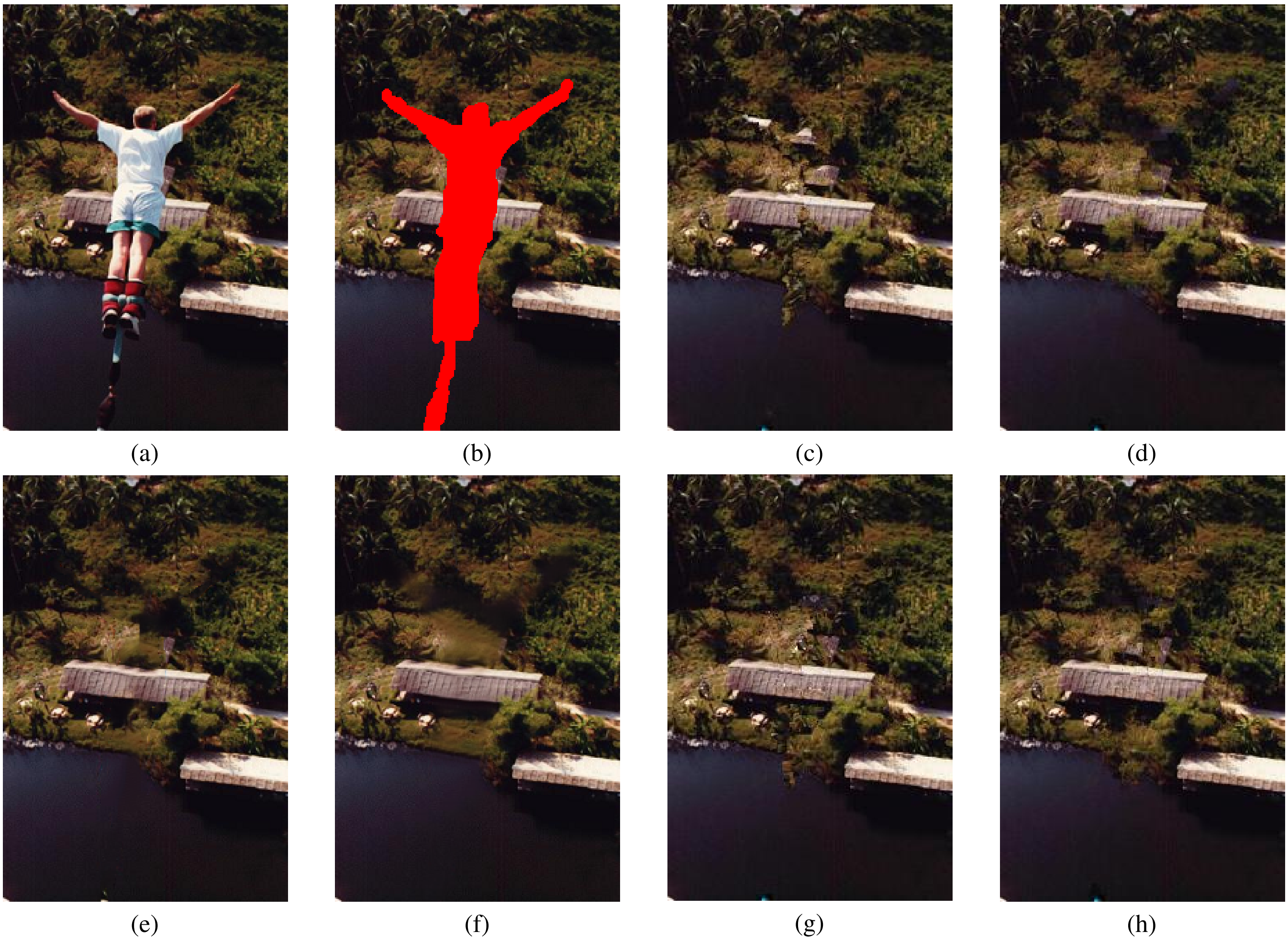

Fig. 3. The experiment on removing an object: (*a*) the original image of bungee jumping man,(*b*) removed region, (*c*) the result of DCT method, (*d*) the result of Criminisi's method, the result of proposed method, (*e*) the result of Wexler's method, (*f*) the result of patchmatch inpainting, (*g*) the result of patch shifting which applying to Criminisi's method, (*h*) the result of proposed method.

shown in Fig. 2 (*c*). PSNR of the result is 35.42dB. Fig. 2(*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. 2(*e*). This method gives the highest PSNR, which is 37.28 dB, due to its nature that try to minimize the global SSD. However, some discontinuity can be noticed as seen on the shoulder of the model. Fig. 2 (*f*) shows the result of patchmatch method which is the fastest method in our experiment. PSNR of Fig. 2 (*f*) is 35.83 dB which is the second worst numerical result of all the techniques we test. Some blur can also be noticed, for example, smoothness at the shoulder of the model. The result of applying patch shifting scheme to Criminisi's method is shown in Fig. 2(*g*). PSNR of the result is 36.24 dB. The reconstructed edge is sharper than Fig. 2 (*c*)-(*f*). However, some discontinuity can be noticed. The result of our proposed method shown in Fig. 2(*h*) has the superb result in visual aspect. In numerical aspect, PSNR of Fig. 2(*h*) is slightly lower than Wexler's method. In this example, 7×7 patch size is used. 10% of the smallest DCT coefficients are ignored and each target patch needs to have known region more than 90% of its area. PSNR of Fig. 2(*d*) is 37.12 dB. The edge at the shoulder of the model is reconstructed perfectly by our method. However, computational time of our method is 2 times higher than the DCT method. Anyway, it is much faster than Wexler's method which usually takes 5-20 time higher than Criminisi's method. Computational time of our proposed method can be reduced by reducing the known region constraint which usually degrades the inpainting result. The worst result our method gives is as same as the result from Criminisi's method.

In Fig. 3, the performance for object removing task is shown. The Bungee jumping man, which is masked as red region in Fig. 3(*b*), on the original image in Fig. 3(*a*) would be removed by various techniques. Fig. 3(*c*)-(*h*) shows the result by Criminisi's method, DCT method, Wexler's method, patch-match method, patch shifting with Criminisi's method and our proposed method respectively. Some discontinuity with strong false edge can be noticed on the result of Criminisi's method in Fig. 3(*c*). The result of smooth edge can be noticed in DCT method as shown in Fig. 3(*d*). In Fig. 3(*e*), Wexler's method gives the result with uneven intensity and a little discontinuity of edge. The lower part of roof top is reconstructed very well as shown in Fig. 3(*f*), however, discontinuity on the upper part of roof top and flatness above the roof top, which makes the image look unnatural, can be noticed. Noise in reconstructed region of 3(*f*) seems to be lower than the known region. In Fig. 3(*g*) the roof top is well reconstructed but the result look noisy comparing with Fig. 3(*f*) and (*h*). The result of our proposed method have a perfectly reconstructed structure (a roof top) as shown in Fig. 3(*h*).

V. CONCLUSION

This paper proposes a novel exemplar-based image inpainting. Patch shifting scheme is introduced and applied to the DCT-based image inpainting. With patch shifting scheme, target patch is more reliable for finding the unknown region because in the case that known region is too small target patch is shifted in the direction that increases the area of known region. The results of our proposed method have a noticeable improvement in numerical and visual quality from the conventional exemplar-based inpainting and DCT-based inpainting. And the visual result is better than most of the well-known exemplar-based inpainting method as shown in section IV. However, computational time is higher than DCT inpainting. In future work, we will study more on how to get a satisfied target patch by less complex approach, for example, enhancing patch priority term to be more sensitive to the data inside the target patch. Patch size is another parameter that we are interested in. Fixed patch size may not field the best performance because different missing region may require different patch size for the best restoration. So framework of adaptive patch size is our promising approach.

REFERENCES

[1] M. Bertalmío, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *SIGGRAPH*, 2000, pp. 417–424.

[2] T. F. Chan and J. Shen, "Local inpainting models and tv inpainting," *SIAM Journal on Applied Mathematics*, vol. 62, no. 3, pp. 1019–1043, 2001.

[3] T. F. Chan and J. Shen, "Non-texture inpainting by curvature-driven diffusions (cdd)," *J. Visual Comm. Image Rep*, vol. 12, pp. 436–449, 2001.

[4] A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in *International Conference on Computer Vision*, 1999, pp. 1033–1038.

[5] M. Bertalmío, L. A. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 882–889, 2003.

[6] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200–1212, 2004.

[7] Tsz-Ho Kwok, Hoi Sheung, and Charlie C.L. Wang, "Fast query for exemplar-based image completion," *IEEE Transactions on Image Processing*, vol. 19, pp. 3106–3115, 2010.

[8] Y. Wexler, E. Shechtman, and M. Irani, "Space-time completion of video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 463–476, 2007.

[9] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman, "PatchMatch: A randomized correspondence algorithm for structural image editing," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 28, no. 3, August 2009.

[10] Sarawut Tae-o-sot and Akinori Nishihara, "Exemplar-based image inpainting with patch shifting scheme," in *17th International Conference on Digital Signal Processing*, July 2011.