The 31st International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2016)

Exemplar-Based Image Inpainting Using Context-Aware Approach

Yusuke Murakami and Koichi Ichige
Department of Electrical and Computer Engineering, Yokohama National University
79-5 Tokiwadai, Hodogaya-ku, Yokohama 240-8501, Japan

Abstract: This paper presents a novel exemplar-based inpainting method using context-aware approach. Exemplar-based methods use texture synthesis techniques and are known to work well when textures are regular or repeatable. One of the conventional methods uses the positional relations between the most similar patches in patch unit, but it cannot achieve the best performance. By applying context-aware approach to it, we propose an high-quality and efficient exemplar-based inpainting method. Performance of the proposed method is evaluated through computer simulation.

1. Introduction

Image inpainting, also known as image completion, is an image processing technique that recovers missing or damaged parts in an image [1]. With recent developments of digital image manipulation, it has become an active research subject of image processing. Applications of image inpainting have many purposes such as old film restoration, art conservation, and digital restoration. Inpainting methods are classified into two main categories: diffusion-based methods and exemplar-based methods.

The diffusion-based methods fill in the missing region by propagating image content from the boundary to the interior of the missing region via based on partial differential equations [1], [2] or variational methods [3]. These methods yield good results when inpainting thin missing region, but these methods tend to give blurred images when inpainting large missing region.

The exemplar-based methods fill in the missing region by copying the best matching texture patches from neighborhood in the known region. In the early 2000s, Bornard *et al.*[4], Drori *et al.*[5], and Criminisi *et al.*[6] proposed independently three exemplar-based inpainting approaches that lay the foundation of important notions of the exemplar-based inpainting approaches. These methods use texture synthesis techniques [7] and are known to work well when textures are regular or repeatable. The first attempt to use exemplar-based techniques for object removal has been reported in [8]. Compared to diffusion-based methods, exemplar-based methods typically produce better results, especially when missing region is larger.

The method in [9] can be regarded as combined the version of their previous methods [10], [11] and the state-of-the-art exemplar-based methods, and improved each key points of the algorithm. This method uses structure tensors [12] to be robust data term for a better selection of pixel candidates to fill in the missing region, and then insuring a global coherence in the reconstruction. It uses the positional relations between the most similar patches in patch unit, but cannot achieve good performance.

In this paper, we propose an efficient and high-quality exemplar-based method with context-aware approach [13]. Context-aware approach divides input image into each structure blocks. It enables to search efficiently for optimal patches by restricting search region. Moreover, it improves the accuracy because this method search only similar structure block for optimal patches.

The rest of this paper is organized as follows. Section 2 introduces conventional method. Section 3 first introduces the Context-aware approach, and then apply it to conventional method. After showing some simulation results in Section 4, we will make concluding remarks in Section 5.

2. Preliminaries

This section briefly prepares the notations and definitions. An image to be inpainted is considered as a function $I : S \to \boldsymbol{R}^3$ where S defines the image domain. Let Ω denote the missing region of the image, and Φ denote the known part of the image, where $\Omega \cup \Phi = S$. In the following, a patch Ψ_p centered on the pixel p is considered as a function $\Psi_p : N_p \to \boldsymbol{R}^3$ where $N_p \in S$ is the square support of Ψ_p. $\Psi_{\hat{p}}$ denotes a patch that matches Ψ_p according to a given metric d:

$$\Psi_{\hat{p}} = \left\{ \Psi_q | \underset{q|N_q \cap \Phi}{\operatorname{argmin}} \, d(\Psi_p, \Psi_q) \right\} \quad (1)$$

The distance d to compare the visual similarity of two patches is the SSD (Sum of Square Difference):

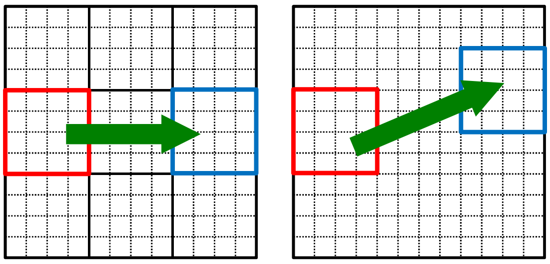
$$d(\Psi_p, \Psi_q) = \sum_{v \in N_p \cap \Phi} ||\Psi_p(v) - \Psi_q(v + p - q)||^2 \quad (2)$$

It is widely used essentially for computational efficiency purpose to compare the visual similarity. If the contexts of patches $\Psi_p(v)$ and Ψ_q are similar, $d(\Psi_p, \Psi_q)$ becomes lower.

3. Conventional Method

The conventional method [9] uses offset statistics of the positional relations between the most similar patches to estimate optimal patches for copying to missing region [14]. The M prevailing offsets can be extracted for each patches in known region, and they are candidates for copying to missing region. Then, optimal patches are determined by computing SSD (Sum of Square Difference) between missing patches and candidate patches. Candidate patches are found by offsets statistics and window search.

This method uses fast and smart window search algorithm. The search window has two types. The first type is the window surrounding the target patch which is commonly used by window search. The other type is the window which has already been used by reconstructing missing region nearby target patch. The search area then consists in several search windows, each centered at the location of the center of the nearby



(a) conventional method [9] (b) ideal

Figure 1. offset statistics

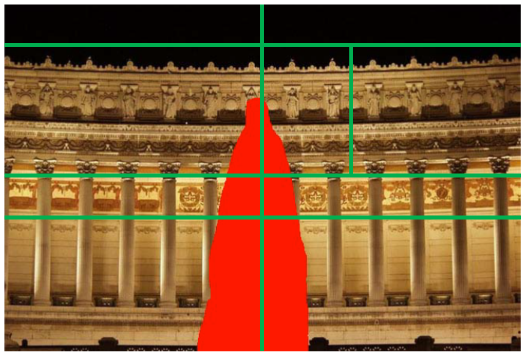


Figure 2. Structure block division

used patch. It can efficiently search for optimal patches by using nearby history informations.

Filling order is determined by computing a priority term consisting of confidence term and data term. Confidence term means a measure of reliable information in the neighborhood of target patch. Data term means the local image structure around the missing region, and it uses structure tensors to be robust.

When patches are fully copied, the block effect artefacts are often produced. It is caused by the copy and paste of patches chunks that do not match perfectly on their common boundaries even with a smart selection scheme. This method blends the pasted patches using tensor model that is very careful about local image structures and textures.

In this method, the most similar patches are found to compute offsets statistics on the whole image that is divided into patches, but it can only compute positional relations in patch unit (as seen is Fig. 1(a)). It has low computational complexity but accuracy of searching the most similar patch is also low. When positional relations are computed on the whole image in a pixel unit (as seen is Fig. 1(b)), the accuracy of searching the most similar patch becomes higher but it takes a long computational time.

4. Proposed Method

In this section, we propose a method of applying the exemplar-based method to the context-aware approach [13]. Context-aware approach divides input image into each structure blocks (as seen is Fig. 2), then the positional relations can be computed in the similar structure blocks. Since search area for the most similar patch is restricted, computational complexity is reduced. Moreover, most patches are similar with target patch in these blocks. It can search the most similar patches from only reliable patches. It causes computing offsets statistics correctly and efficiently, then it means improving the accuracy and reducing the computational time.

The proposed procedure can be briefly summarized as the following steps.

1. Apply Gabor filters to the input image and apply K-means clustering to the magnitudes of complex responses. Then, each pixel p is assigned to one of the K textons [15], and $T(p)$ will denote this pixel-to-texton mapping.
2. Evaluate the block's homogeneity under increasing a block size and changing a increasing direction horizontal and vertical alternately. If the block's homogeneity is higher than threshold, the image is divided into two blocks.
3. Compute the offset statistics of the positional relations between the most similar patches by searching similar structure blocks in a pixel unit.
4. Determine the filling order by computing a priority term consisting of confidence term and data term.
5. Estimate optimal patches and copy to missing region until missing region is filled completely. Then, optimal patches are estimated from candidate patches. If the missing region still exists, return to 4.
6. Blend the pasted patches to reduce the block effect artefacts caused by the copy and paste of patches chunks that do not match perfectly on their common boundaries.

Steps 1 to 3 are the difference between proposed method and conventional method [9]. Steps 1, 2 are newly added to the conventional method, and Step 3 modifies the changed searching interval from patch unit to pixel unit.

To compute the block homogeneity, contextual descriptors as some characterization of spatial content and textures within blocks are used. The contextual descriptor $c_n^{(\ell)}$ of the ℓ-th block B_ℓ becomes

$$c_n^{(\ell)} = \frac{1}{|B_\ell \cap \Phi|} \sum_{p \in B_\ell \cap \Phi} \xi[\mathbf{T}(p) = n], \tag{3}$$

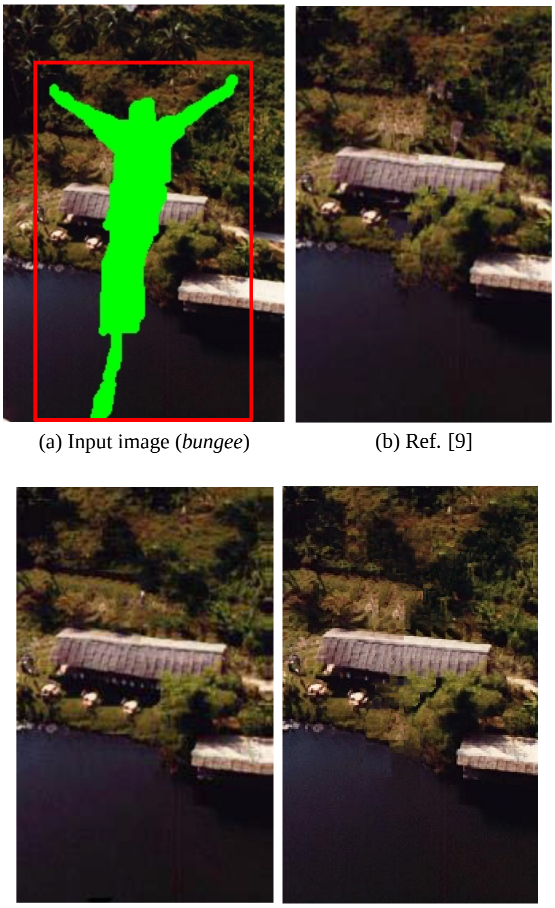
where $|\cdot|$ denotes the cardinality of the set and ξ returns one if its argument is true and zero otherwise. $T(p)$ takes one of the K possible values, for K textons, i.e., $T(p) = n$, $n = 1, ..., K$. Then, the block homogeneity $H^{(\ell,m)}$ between ℓ-th block and m-th block can be written as

$$H^{(\ell,m)} = \frac{1}{2} \sum_{n=1}^{K} \frac{\left(c_n^{(\ell)} - c_n^{(m)}\right)^2}{\left(c_n^{(\ell)} + c_n^{(m)}\right)}. \tag{4}$$

The value of $H^{(\ell,m)}$ becomes larger, if the contexts of blocks B_ℓ and B_m are not similar.

5. Simulation

In this section, we evaluate the proposed method by comparison with conventional methods. The patch size of our simulation is changed for each image.



(a) Input image (*bungee*) (b) Ref. [9]

(c) Ref. [13] (d) Proposed

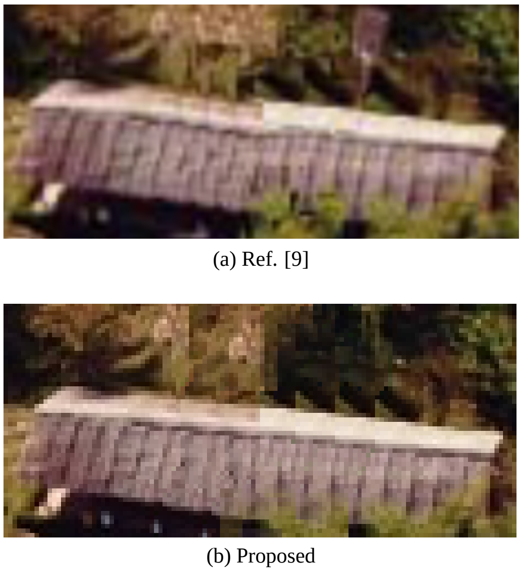
Figure 3. Comparisons of the original and inpainted results for the image *bungee*

5.1 Comparison with conventional methods

Figure 3 shows the comparison of our inpainting result with the conventional methods, and Fig. 4 shows the enlarged images of Fig. 3(a) and (b). Image size of *bungee* is 206×308. The patch size of our simulation is the same size as Ref. [9] (17×17). In comparison with Fig. 3(b), our result is recovered more naturally (especially the roof and the region over the building). On the other hand, Fig. 3(c) looks better than our result, but Ref. [13] needs to iterate at most 10 times, and it takes twice the computational time at each iteration. Figure 5 shows the comparison of two other images. Image size of *yard* is 440×330, and *cow* is 600×400. The patch size of our simulations are 17×17 (*yard*) and 21×21 (*cow*). As shown in Fig. 5, our proposed method exhibits visually equivalent or better than conventional method.

5.2 Comparison of the inpainted results with different patch size

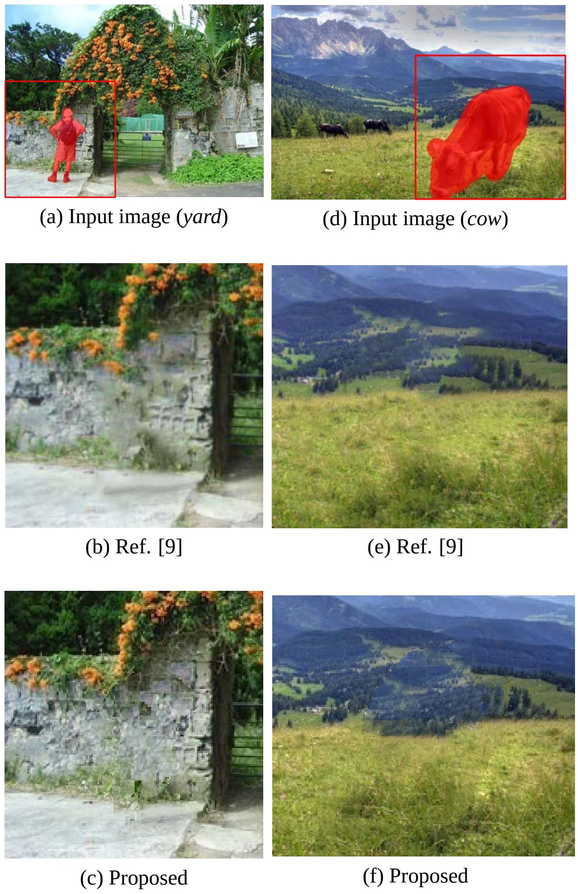
Optimal values of parameters are needed to tune for each image. If optimal parameters is not be able to set, the inpainting result may degrade than the best results. Figure 6 shows the comparison different patch size results for *bungee* and *yard*. As shown in Fig. 5, *bungee* is the best result with 17×17 patch size, and *yard* is the best result with 21×21 patch size. Therefore, the optimal patch size is changed at each image. Optimal patch size depends on the size, resolution, and the content of the image. We will consider the way to determi-



(a) Ref. [9]

(b) Proposed

Figure 4. Enlarged results (Fig. 3(a), (b))

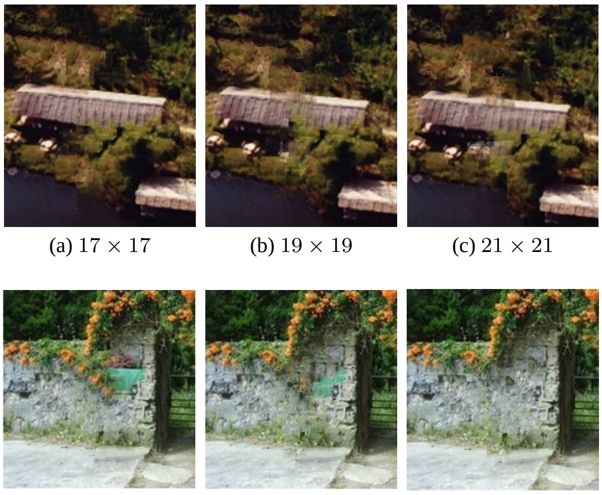


(a) Input image (*yard*) (d) Input image (*cow*)

(b) Ref. [9] (e) Ref. [9]

(c) Proposed (f) Proposed

Figure 5. Comparisons of the original and inpainted results for the images *yard* and *cow*



(a) 17×17 (b) 19×19 (c) 21×21

(d) 17×17 (e) 19×19 (f) 21×21

Figure 6. Comparison of the inpainted images with different patch size

nate the optimal patch size using those factors as one of future studies.

6. Concluding Remarks

This paper presents a novel exemplar-based inpainting method with context-aware approach. It can find the most similar patches efficiently and correctly. We confirmed through computer simulation results that the proposed method worked superior than the conventional methods, and the optimal patch size was different depending on the characteristics of the original images.

References

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballseter, “Image inpainting,” *Proc. 27th Annu. Conf. SIGGRAPH*, pp.417-424, July 2000.

[2] D. Tschumperlé and R. Deriche, “Vector-valued image regularization with PDEs: a common framework for different applications,” *IEEE Trans. on Pattern Analysis and Machine Intell.*, vol. 27, no. 4, pp.506-517, Mar. 2005.

[3] T. Chan and J. Shen, “Variational restoration of nonflat image features: Models and algorithms,” *SIAM J. Appl. Math.*, vol. 61, no. 4, pp.1338-1361, 2001.

[4] A. Criminisi, P. Pérez and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” *IEEE Trans. Image Process.*, vol. 13, no. 9, pp.1200-1212, Aug. 2004.

[5] I. Drori, D. Cohen-Or and H. Yeshurun, “Fragment-based image completion,” *Proc. ACM SIGGRAPH*, vol. 22, no. 3, pp.303-312, July 2003.

[6] C. Barnes, E. Shechtman, A. Finkelstein and D. B. Goldman, “PatchMatch: A randomized correspondence algorithm for structural image editing,” *Proc. ACM SIGGRAPH*, vol. 24, pp.1-11, Aug. 2009.

[7] A. A. Efros and T. K. Leung, “Texture synthesis by non-parametric sampling,” *Proc. 7th IEEE Int. Conf. Compt. Vis.*, vol. 2, pp.1033-1038, Sep. 1999.

[8] P. Harrison, “A non-hierarchical procedure for re-synthesis of complex texture,” *Proc. Int. Conf. Central Europe Comp. Graphics, Visua. and Comp. Vision, Plzen, Czech Republic*, pp.190-197, Feb. 2011.

[9] P. Buyssens, M. Daisy, D. Tschumperlé and O.Lézoray, “Exemplar-based inpainting: Technical review and new heuristics for better geometric reconstructions,” *IEEE Trans. Image Process.*, vol. 24, no. 6, pp.1809-1824, June. 2015.

[10] M. Daisy, D. Tschumperlé, and O. Lézoray, “A fast spatial patch blending algorithm for artefact reduction in pattern-based image inpainting,” *Proc. SIGGRAPH Asia Tech. Briefs*, pp. 1-4, Nov. 2013.

[11] M. Daisy, P. Buyssens, D. Tschumperlé, and O. Lézoray, “A smarter exemplar-based inpainting algorithm using local and global heuristics for more geometric coherence,” *Proc. IEEE Int. Conf. Image Process.*, pp. 4622-4626, Oct. 2014.

[12] S. D. Zenzo, “note on the gradient of a multi-image,” *Comput. Vis. Graph. Image Process.*, vol. 33, no. 1, pp. 116-125, 1986.

[13] T. Ružić and A. Pižzurica, “Context-aware patch-based image inpainting using Markov random field modeling,” *IEEE Trans. Image Process.*, vol. 24, no. 1, pp.444-456, Jan. 2015.

[14] K. He and J. Sun, “Statistics of patch offsets for image completion,” *Proc. 12th Eur. Conf. Comput. Vis.*, vol. 7573, pp. 16-29, 2012.

[15] J. Malik, S. Belongie, T. Leung and J. Shi, “Contour and texture analysis for image segmentation,” *Int. J. Comput. Vis.*, vol. 43, no. 1, pp.7-27, Jun. 2001.