

Exemplar-based Image Inpainting using Structure Tensor^{*}

Liu Kui^{1,2} Tan Jieqing¹ Su Benyue²

¹School of Computer and Information, Hefei University of Technology, Hefei 230009, China

²School of Computer and Information, Anqing Normal University, Anqing 246011, China
liukui@aqtc.edu.cn

Abstract – Exemplar-based image inpainting techniques face two main problems. One is the decision of filling-in order which has a strong impact on image inpainting effect. Another problem is greedy search process while searching for the best exemplar patch. This paper presents a robust method for exemplar-based image inpainting based on structure tensor. The novel priority function is redefined. Experimentation results show that our image inpainting method improves the quality of image inpainting compared with the conventional exemplar-based image completion algorithms.

Index Terms – Image inpainting, texture synthesis, exemplar, structure tensor, structural similarity (SSIM)

I. Introduction

Image inpainting technique remains a longstanding challenge in image processing and computer vision, which aims to use the known image information to restore or fill damage or missing parts of an image or a video in an undetectable way for the ordinary observers who aren't familiar with the original images. Nowadays, it has been widely applied in restoration of scratched old photos and films, heritage conservation, removal of objects in images, virtual reality, and so on.

In recent years, Lots of efforts have been made to develop efficient image inpainting methods. Generally speaking, existing image inpainting approaches can be classified into two categories: diffusion-based techniques and texture-oriented techniques.

The first category is the diffusion-based approach first proposed by Bertalmio et al. [1] in 2000, which is founded on the theory of partial differential equation (PDE). These methods aim at filling the missing parts by diffusing the image information from the known region into the missing region along the isophote. Inspired by Bertalmio's work, Chan and Shen proposed the total variational (TV) inpainting model [2]. The TV model employs anisotropic diffusion based on the contrast of the isophotes and is beneficial to preserving the structure during the reconstruction. Later, they developed curvature-driven diffusion model (CDD) [3] which extended TV model by taking geometric information of isophotes into account when defining the strength of the diffusion process. Although diffusion-based inpainting algorithms have achieved convincingly excellent results for completing small, non-textured target region, they tend to introduce over-smooth effect or stair-case effect in the textured region or larger missing region [4]. The second category of approaches is the texture-oriented inpainting algorithms, which are originated

from texture synthesis in [5]. Efros and Leung first proposed the exemplar-based texture synthesis model [6], and then Bornard introduced this model into the geometrical natural image completion [7]. Cheng et al. proposed a robust image inpainting approach to determine the filling order by the structure priority value and to find the best patches in the dynamic searching range in the source region [8]. Wong proposed an exemplar-based nonlocal means approach for the image inpainting [9], in which the image patch was determined by a set of candidate patches in the known region instead of a single best match patch. Up to now, the most famous exemplar-based inpainting method [10] is proposed by Criminisi et al, in which the inpainting order is decided by a predefined priority function to ensure that the linear structures will propagate before texture filling to preserve the connectivity of object boundaries. Wu proposed a new exemplar-based inpainting algorithm, in which a cross-isophotes patch priority term was defined by anisotropic diffusion [11]. The experimental results of exemplar-based inpainting algorithms show that they are very effective in reducing the blurring artifacts and suitable for both the small and large image gaps [12]. However, a major defect of their approaches is their greedy search for the most similar patch in the source region, which can cause visual inconsistencies. In addition, the exemplar-based inpainting algorithms have some drawbacks in calculating the priority. Especially, when the strong structure exists in the texture image, structure discontinuous phenomenon appears.

In this paper, we present a robust image inpainting method based on structure tensors because the structure tensor provides a more powerful descriptor of local pattern images better than a simple gradient, which has been proved in [13].

The rest of this paper is organized as follows. Section 2 briefly introduces the classical Criminisi algorithm. In Section 3, structure tensor is presented. The improved algorithm based on structure tensor is presented in the section 4. Section 5 shows some experimental results. Finally, conclusions are drawn in section 6.

II. Exemplar-based image inpainting model

A. Review of the exemplar-based algorithm

For the convenience of explanation, the notations adopted in the paper are similar to that used in Criminisi's algorithm. The region to be filled, also called the target region, is denoted

^{*} This work was supported by the NSFC-Guangdong Joint Foundation Key Project (Grant No. U1135003), the National Natural Science Foundation of China (Grant No.60773043, 61070227) and the Foundation for Key Program of Ministry of Education of China (No. 309017).

by Ω , and the source region is represented as Φ , which provides samples used in the reconstruction process. Let $\delta\Omega$ be the boundary between Ω and Φ . So, the entire image I is composed of two regions: the source region Φ and the target region $\Omega(\Omega=I-\Phi)$, as shown in figure 1.

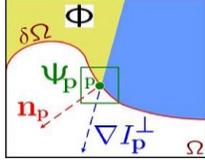


Fig. 1 Notation diagram [10]

Generally speaking, an exemplar-based image inpainting algorithm includes four main steps as follows:

Step 1 Identify the target region. The target regions are extracted and the contours are identified manually.

Step 2 Compute filling priorities and select the patch with the highest priority. Each patch Ψ_p centered at the pixel p ($p \in \delta\Omega$) has a inpainting priority, which is defined to be the product of Confidence term $C(p)$ and Data term $D(p)$.

$$P(p) = C(p)D(p) \quad (1)$$

The confidence term $C(p)$ is thought of the number of known information in the patch Ψ_p and is defined as:

$$C(p) = \frac{\sum_{q \in \Psi_p \cap \Phi} C(q)}{|\Psi_p|} \quad (2)$$

Where $|\Psi_p|$ is the area of patch Ψ_p . During initialization, $C(p)$ is often set to 0 for all pixels in the target region Ω , and to 1 for known pixels in the source region Φ .

The data term $D(p)$, which encourages propagating linear structures, is a function of the strength of isophotes hitting the front $\delta\Omega$ at each iteration and is defined below:

$$D(p) = \frac{|\nabla I_p^\perp \cdot n_p|}{\alpha} \quad (3)$$

Where α is a normalization factor (e.g., $\alpha = 255$ for a grey-level image), I_p^\perp is isophote at pixel p and n_p is a unit vector orthogonal to the contour $\delta\Omega$ at the pixel p . Then find the patch $\Psi_{\hat{p}}$ with the highest priority.

Step 3 Search the most similar block $\Psi_{\hat{q}}$ with $\Psi_{\hat{p}}$ from the source region Φ . Copy the pixels from $\Psi_{\hat{q}}$ to $\Psi_{\hat{p}}$.

Step 4 Update confidence values and the contour of the the target region.

The algorithm iterates the above four steps until all pixels in the target Ω have been filled.

B. Analysis of the exemplar-based algorithm

Filling order is crucial for exemplar-based inpainting algorithms. In Criminisi's algorithm, the filling priority is the product of the data term $D(p)$ and the confidence term $C(p)$, which can make the algorithm retain not only structure information but also texture information. The data term contains isophote structure information. The confidence term measures the reliable information around the pixel p needed inpainted. But the result of formula (1) isn't stable to unexpected noise and extreme values. In fact, when the value of $D(p)$ is large, linear structure information should earlier be completed than texture information. However, when the patch centered at the pixel p has too few pixels of the source region, the corresponding data term becomes not reliable. In experiments, we find that the value of confidence term drops too fast to zero while the value of data term is quite steady. In the paper, we first regularize the confidence term, and then use the weight sum of confidence term and data term to generate the dependable priority.

In addition, only considering the direction of isophote when computing the value of the data term, it easily leads to estimate inaccurate for the inpainting order, which also make the reconstructed image authentic. To overcome this flaw, we introduce structure tensor in calculation of the data term $D(p)$. Since the structure tensor contains not only intensity information in a local area but also the predominant directions of the gradient in a specified neighborhood of one pixel and the degree to which these directions are coherent.

III. Structure tensor

Structure tensor, also named the second-moment matrix, is a reliable tool in image processing and computer vision. Structure tensor contains the information on orientation and magnitude of the image structures and the information which measures the homogeneity of orientations within the neighborhood of a pixel. It has been used in many application fields such as optical flow estimation, corner detection, and texture analysis.

Consider the 2-dimensional image domain $\Omega \subset \mathfrak{R}^2$, and let an image $I(x)$ be represented by a bounded mapping $I: \Omega \rightarrow \mathfrak{R}$. The gradient vector value of one pixel is defined as $(\nabla I) = (I_x, I_y)^T$, and the structure tensor is defined as:

$$J = g_\rho \otimes J_0 = \begin{bmatrix} g_\rho \otimes I_x^2 & g_\rho \otimes I_x I_y \\ g_\rho \otimes I_x I_y & g_\rho \otimes I_y^2 \end{bmatrix} \quad (4)$$

Where the operator \otimes is the convolution operator and g_ρ is the Gaussian function. The orthogonal eigenvectors v_1 and v_2 of the structure tensor J are computed as follows:

$$\begin{cases} v_1 = (2J_{12}, J_{22} - J_{11} + \sqrt{(J_{22} - J_{11})^2 + 4J_{12}^2})^T \\ v_2 \perp v_1 \end{cases} \quad (5)$$

The corresponding eigenvalues u_1 and u_2 of the structure tensor J are calculated by Eq. (6).

$$\begin{cases} u_1 = \frac{1}{2}(J_{11} + J_{22} + \sqrt{(J_{22} - J_{11})^2 + 4J_{12}^2}) \\ u_2 = \frac{1}{2}(J_{11} + J_{22} - \sqrt{(J_{22} - J_{11})^2 + 4J_{12}^2}) \end{cases} \quad (6)$$

We know that eigenvalues u_1 and u_2 can reflect the shape characteristic of the local position [14]. Both the eigenvalues are small, namely $u_1 \approx u_2 \approx 0$, which means the pixel is located in the smooth area of the image; One eigenvalue is large and the other one is small, namely $u_1 \gg u_2 \approx 0$, which means the pixel is located in the edges; Both the eigenvalues are rather large, namely $u_1 \geq u_2 \gg 0$, which means that the pixel is the corner.

By analyzing the feature of structure tensor, we can obtain pixel information in a local area but also the predominant directions of the gradient in a specified neighborhood of one pixel and the degree to which these directions are coherent by structure tensor. These features of the structure tensor are beneficial to determining the filling order of the target region when the structure tensor is used to compute $D(p)$.

IV. Improved algorithm

We find that the value of $C(p)$ tends to zero with the proceed of the original algorithm, so a regularized confidence term shown in [8] takes the place of the original confidence term.

$$C_N(p) = (1 - \omega)C(p) + \omega \quad 0 < \omega \leq 1 \quad (7)$$

where ω is regularizing factor. So the value of the new confidence term is regularized to $[\omega, 1]$.

The calculation function of the data item $D(p)$ is the most critical point of the algorithm. The accurate estimation of the edge direction determines the value of the data item directly. If there is a lot of texture information at the edge of the image, it may lead to inaccurate estimation of the structure edge. So we introduce structure tensor to improve the calculation method of $D(p)$.

$$D(p) = \text{div}(J\nabla I(x, y)) \quad (8)$$

where div is divergence operator. After the improvement, the restoration order of the target patch is no longer only determined by ∇I_p^\perp .

The priority function is defined as the weight sum of regularized confidence term $C(p)$ and new data term $D(p)$.

$$P(p) = \alpha C_N(p) + (1 - \alpha)D(p) \quad (9)$$

where α is adjustment coefficient, satisfying $0 < \alpha < 1$. The value of α is not adaptive to the information of the image while it is decided by the many tests.

The improved image inpainting algorithm is given as follows:

Initially, select the target patch to be inpainted manually, extract the contour of the patch, and then repeat the following steps until the inpainting is completed.

- Step 1 : Determine the edge of the patch to be inpainted $\delta\Omega$;
- Step 2 : Calculate $C_N(p)$, $D(p)$ and $P(p)$;
- Step 3 : Find the inpainting block $\Psi_{\hat{p}}$ with the highest priority;
- Step 4 : Search the image block $\Psi_{\hat{q}}$ which is best matched with the current block $\Psi_{\hat{p}}$ to be inpainted in the source patch of the image;
- Step 5 : Update the confidence term $C(p)$ which has been filled;

Repeat the above steps until the inpainting is completed.

V. EXPERIMENTAL RESULTS

To demonstrate the validity of our inpainting approach, we test the proposed algorithm on a variety of natural images or synthetic images and compare the experimental result of our approach with that of the Criminisi algorithm [10] and the algorithm in [11]. Evaluation of the efficiency of the image inpainting algorithm is generally done with subjective assessment because there is no original undamaged image in reality. But we often generate damaged images from some known originals, and then the quantitative evaluation is performed by comparing PSNR (the Peak Signal-to-Noise Ratio) between the restored version and the original ones, for example Fig.2 and 3. Generally the higher the PSNR value, the larger the similarity of the repaired image to the original. The equations to calculate PSNR are given as follows.

$$MSE = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [u(x, y) - u_0(x, y)]^2 \quad (10)$$

$$PSNR = 10 \lg \frac{255^2}{MSE} \quad (11)$$

However, PSNR sometimes disaccords with human visual judgments. So, SSIM (Structural Similarity) is used as well because the SSIM criterion is closer to the human vision system [15].

In all experiments, the size of patches is set to 7 by 7, 9 by 9 or 15 by 15 for all images. All the experiments are run on the computer with the configuration of 2.1GHz processor and 2GB RAM.

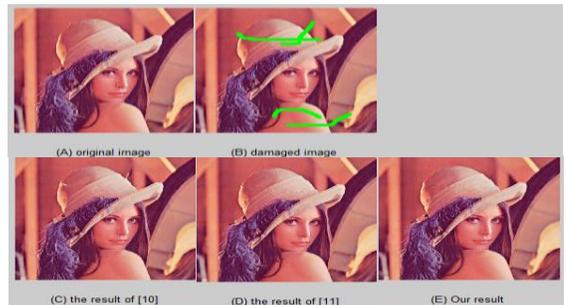


Fig.2 scratch restoration of Lena

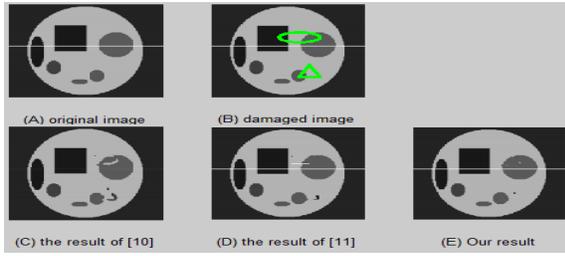


Fig.3: scratch restoration of synthetic image

Fig.2 and 3 present the inpainted results in the applications of scratch restoration. The (A) column of Fig2 and 3 are the original images, and Fig. 2 and 3 (B) are the scratch images. Seen from Fig.2, the result of Criminisi's algorithm has artificial trace in the region of the lena's cap and her hair. Our result is a little better than that of [11] in the lena's hair. Fig.3 further proves that our method is better than papers [10] and [11].

Table 1 PSNRs and SSIMs of different images

image	PSNR			SSIM		
	Proposed algorithm	Algorithm of[10]	Algorithm of[11]	Proposed algorithm	Algorithm of[10]	Algorithm of[11]
Lena	53.8855	52.5617	53.8495	0.9850	0.9825	0.9848
Synthetic image	55.9038	50.0549	53.6863	0.9854	0.9691	0.9849

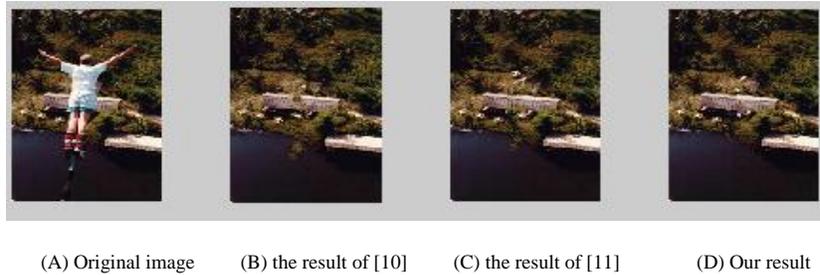


Fig. 4: The removal of the big object

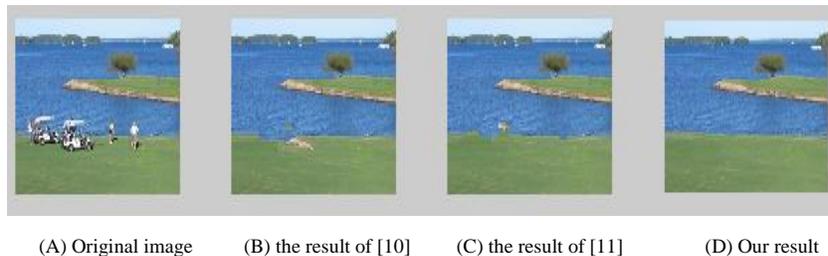


Fig. 5 The removal of the Multi-target objects

Table 1 shows the comparison of the PSNR and SSIM for our method with the other two image inpainting techniques in papers [10] and [11].

Figure 4 shows the removal of the big object. Fig. 4 (A) is the original image. Fig.4 (B) and Fig.4(C) are respectively the completed images using the exemplar-based model in [10] and the exemplar-based model in [11]. In Fig.4 (B), there are many error texture patterns filled and the completed result of the riverbank and roof is not verisimilar. Fig. 4 (D) is the inpainting result of our model. The visual perception of the image in Fig.4 (D) is better, because the riverbank is completed with a plausible riverside and the structure of roof

is preserved. But the computation time of our algorithm is a little longer than that of [10] or [11]. The result of our inpainting is better than papers [10] and [11].

Figure 5 shows the multi-objects removal, where Fig.5 (A) is the original image, Fig.5(B) represents the completed image using the exemplar-based model in [10], Fig.5 (C) shows the completed image using the algorithm in [11], Fig.5 (D) shows the inpainting result of algorithm using our model. The images in Fig.5 (B) and Fig.5 (C) are not authentic. However, the inpainting result of our algorithm in the paper is natural.

VI. Conclusion

Since there are some defections in calculating the priority in the criminisi algorithm, such as the inadequate calculation of the priority lead to the unreliable inpainting order. In this paper a robust exemplar-based image inpainting algorithm is proposed. The algorithm uses a novel priority function to determine the inpainting order. Structure tensor is adopted in the priority function. Structure tensor contains the transformation direction of the image and the size of transformation along these directions. Therefore, a patch along a strong geometric structure has a higher inpainting priority and would be completed prior to other patches to preserve the linear structure. Additionally, as seen in our experiments, the weighted sum of the priority can achieve better results than other traditional exemplar-based inpainting model. Our future work is that we plan to investigate the similarity between the source patch and the object patch because only considering the color distance is not reasonable.

References

- [1] Bertalmio M., Sapiro G., Caselles V., Ballester C. Image inpainting, in Proceedings of SIGGRAPH, New Orleans, LA, 2000, July, pp: 417–424.
- [2] Chan T. F., Shen J. Mathematical models of local non-texture inpaintings. *SIAM Journal on Applied Mathematics*, 2001, 62(3): 1019-1043.
- [3] Chan T. F., Shen Jian-hong. Non-texture Inpainting by Curvature-driven Diffusions(CDD). *J. Visual Comm. Image Rep.*, 2001, 4(12): 436-449.
- [4] Chan T. F., Kang S.-H., Shen J. Euler's Elastica and Curvature-based Image Inpaintings, CAM Report 01-12, Department of Mathematics, UCLA, Los Angeles, CA, 2001.
- [5] Efros A. A. and Leung T., Texture synthesis by non parametric sampling. In: Proc. International Conference in Computer Vision, Greece, 1999, 1033-1038
- [6] Efros A. A. and Freeman W. F., Image quilting for texture synthesis and transfer, in Proceedings of ACM Conference in Computer Graphics, 2001, pp. 341-346.
- [7] Bornard R.E., Laborelli L., and Chenot J.H., Missing data correction in still images and image sequences, in Proceedings of ACM Multimedia, 2002, pp. 355-361.
- [8] Cheng W H, Hsieh C W, Lin S K, et al. Robust algorithm for exemplar-based image inpainting, Proceedings of the International Conference on Computer Graphics, Imaging and Visualization. 2005.
- [9] Wong A. and Orchard J., A nonlocal-means approach to exemplar based inpainting, presented at the IEEE Int. Conf. Image Processing, 2008.
- [10] Criminisi A., Perez P., Toyama K., Patch filling and Object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, 2004, 13(9): 1200-1212.
- [11] Wu Ji-ying, Ruan Qiu-qi, A novel exemplar-based image completion model. *Journal of Information Science and Engineering*, 2009(25): 481–497.
- [12] Chang I C, Hsu C W. 15 Image Inpainting Using an Enhanced Exemplar-Based Algorithm. *Multimedia Security and Stenography*, 2012: 275.
- [13] Weickert J. Theoretical foundations of anisotropic diffusion in image processing. *Computing Supplement*, 1996, 11: 221- 236
- [14] Weickert J. Coherence-enhancing diffusion filtering. *International Journal of Computer Vision*, 1999, 31(2) : 111-127.
- [15] Zhou W, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.