ARTICLE IN PRESS
JID: NEUCOM [m5G;June 7, 2017;9:52]
Neurocomputing 000 (2017) 1–7
Contents lists available at ScienceDirect
Neurocomputing
journal homepage: www.elsevier.com/locate/neucom

Exemplar-based image inpainting using structure consistent patch matching

Haixia Wang[a], Li Jiang[a], Ronghua Liang[a,*], Xiao-Xin Li[b]

[a] College of Information Engineering, Zhejiang University of Technology, 310023 Hangzhou, China
[b] College of Computer Science and Technology, Zhejiang University of Technology, 310023 Hangzhou, China

ARTICLE INFO

Article history:
Received 30 November 2015
Revised 11 June 2016
Accepted 21 August 2016
Available online xxx

Keywords:
Image inpainting
Exemplar-based
Patch matching
Fast Fourier transform

ABSTRACT

Image inpainting restores lost or deteriorated parts of images according to the information of known regions. Criminisi has proposed an effective exemplar-based inpainting method, which has the advantages of both texture synthesis and diffusion-based inpainting. Yet, it has its own flaws of fast priority dropping and visual inconsistency. In this paper, we propose a space varying updating strategy for the confidence term and a matching confidence term to improve the filling priority estimation. We propose structure consistent patch matching to take the distribution of source and target patch differences into account. Fast Fourier transform is adapted for full image searching to achieve better and faster matching results. Experimental results are given to demonstrate the improvements made by our proposed method.

1. Introduction

The word image inpainting can be traced back to Renaissance when artists repaired missing or damaged parts of rare artworks by hands. With the development of modern computers and digital image processing, automatic image inpainting has been realized without the requirements of artificial knowledge and professional skills. Image inpainting has broad applications in image interpolation, photo restoration, zooming and super-resolution, occlusion removal, image compression, etc. thus attracts wide research attentions [1–3].

Image inpainting is an ill-pose problem since there is no well-defined unique solution. The pixels in the unknown regions are assumed to have same statistical properties or geometrical structures with some pixels in the known regions, based on which they are estimated. There are mainly two categories of image inpainting methods. The first category is the diffusion-based inpainting, where parametric models are established by partial differential equations to propagate the local structures from known regions to unknown regions. It was first introduced by Bertalmio et al. and realized as the Bertalmio–Sapiro–Caselles–Ballester method [4]. This method simulates the repairing process of professional artists by propagating image information along the isophote. Inspired by this idea, Chan and Shen proposed the Total Variation (TV) inpainting model which can maintain edge information and perform denoising at the same time using anisotropic diffusion [5]. Later, Chan and Shen improved the method by replacing the TV model with a curvature driven diffusion model [6]. The diffusion-based inpainting methods can achieve superior performance when applied to images composed of structure information or when the target region is small. However, when large target regions or natural images composed of both texture and structure information are targeted, diffusion-based inpainting methods introduce blur to the restored images.

The other category of image inpainting methods is regarded as the exemplar-based inpainting. The idea is inspired by the seminal work on texture synthesis [7,8], which grows a large texture image from a given small seed image so that the texture image has a similar visual appearance as the seed image [7]. Similar to region-growing, exemplar-based inpainting restores the unknown regions one pixel or one patch at a time while maintaining coherence with nearby pixels. Ashikhmin searches the best match for an unknown pixel only among the set of shifted candidates from the correspondents of the neighbors [9]. A certain coherence is further imposed in the mapping function to improve the visual qualities of the synthesis results [10]. Efros and Leung proposed an inpainting method that fills the unknown patches instead of unknown pixels one at a time that greatly improves the inpainting speeds [7]. Criminisi et al. presented an exemplar-based image inpainting method dealing both structure and texture information. It adopts the non-parametric sampling concept and sets the known regions of the image as source exemplars. Patches with more edge information are assigned with higher priorities [11]. Recently, the scope of the source exemplars is extended to a database of images instead of a

* Corresponding author.
E-mail addresses: hxwang@zjut.edu.cn (H. Wang), jl@zjut.edu.cn (L. Jiang), rhliang@zjut.edu.cn (R. Liang), mordekai@zjut.edu.cn (X.-X. Li).

http://dx.doi.org/10.1016/j.neucom.2016.08.149
0925-2312/© 2017 Published by Elsevier B.V.

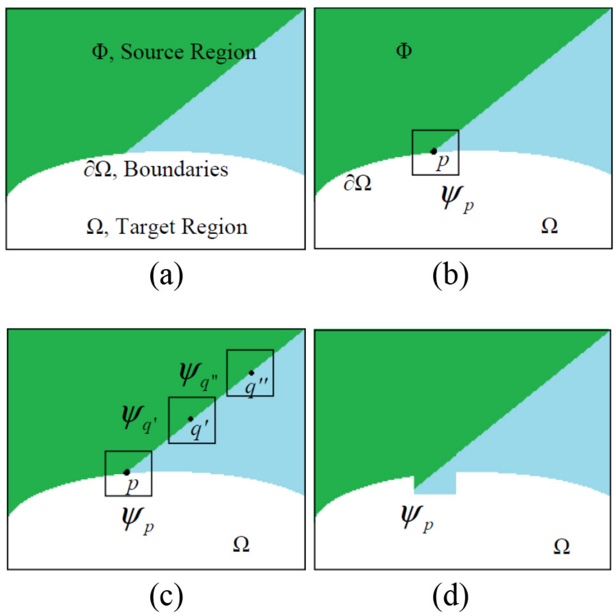

Fig. 1. Structure propagation by exemplar-based inpainting. (a) Source and target regions; (b) target patch; (c) target patch and source patches; (d) filling result of one patch.

single image, which is more effective but also more time consuming [12].

The Criminisi method has achieved great success due to its simple and effective matching criterion and its priority function considering both pixel confidence and structure information, which however have their own flaws. The confidence term in the priority function has fast dropping effect which decreases the inpainting performance [13]. The criterion used for patch matching cannot fully reflect the similarity of two patches. A weighted Bhattacharya distance has been proposed to cope with this problem (denoted as the CBD method) which however only values the overall intensity distribution but not the spatial distribution [14].

We propose an exemplar-based image inpainting using structure consistent patch matching method, which has the following merits. First, a space varying updating strategy for the confidence term and a matching confidence term are proposed to reduce the fast dropping effect and improve the priority estimation, which is critical to the final inpainting results. Second, we propose a structure consistent patch matching to take the distribution of patch differences into account. In the last, fast Fourier transform (FFT) is adapted for full image searching to achieve better and faster matching results.

The remainder of this paper is organized as follows. In Section 2, we briefly introduce the Criminisi method. In Section 3, we discuss the limitations of Criminisi method and propose our exemplar-based image inpainting using structure consistent patch matching. Experimental results and comparisons are given in Section 4 to demonstrate the performance of the proposed method. A conclusion is drawn in Section 5.

2. Criminisi method

Criminisi method has the advantages of both diffusion-based inpainting and texture synthesis. It can restore structure and texture information simultaneously [11] and achieves good performance on images with large lost regions and images composed of textures and structures.

For an image I as shown in Fig. 1(a), Φ indicates the source region which is known, Ω indicates the target region which needs to be filled, and $\partial\Omega$ is the boundary of the target region connecting Φ where inpainting starts. For points in Ω, the points filled earlier can affect the points filled afterwards. The processing priorities of these points are important to final result. A priority function is thus established for every $p \in \partial\Omega$ based on a patch ψ_p surrounding point p as shown in Fig. 1(b). With center point located at $\partial\Omega$, ψ_p has points partially belonging to Φ and partially belonging to Ω. Criminisi method considers that a point has higher priority when the known points in ψ_p have higher confidence and present more structure information than the others. The priority $P(p)$ is defined as

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

The confidence term $C(p)$ measures the amount of reliable information surrounding the point p, which is estimated as

$$C(p) = \frac{\sum_{t\in\psi_p\cap\Phi} C(t)}{|\psi_p|}, \tag{2}$$

where t indicates the coordinates of points belonging to both ψ_p and Φ; the confidences of points in the source region are initialized to 1; $|\psi_p|$ is the total number of points in the target patch ψ_p. More known points in ψ_p each with higher confidence results in higher $C(p)$ value. Meanwhile, the data term $D(p)$ represents the strength of the isophote hitting the boundary $\partial\Omega$ as

$$D(p) = \frac{\left|\nabla I_p^{\perp} \cdot \vec{n}_p\right|}{\alpha}, \tag{3}$$

where α is a normalization factor ($\alpha = 255$ for gray-level images), $\vec{n}_p$ is a unit vector orthogonal to the boundary $\partial\Omega$ at point p, and $\perp$ denotes the orthogonal operator. Points where structure information located have higher priorities.

After a pixel $\widehat{p} \in \partial\Omega$ with highest priority is selected, unknown points in the target patch $\psi_{\widehat{p}}$ need to be filled with the most similar source patch from source region Φ. As Fig. 1(c) shows, both patches $\psi_{q'}$ and $\psi_{q''}$ are seen to be similar to the target patch $\psi_{\widehat{p}}$. How to select the correct source patch, or how to perform the patch matching, is most important to the image inpainting quality. Criminisi method uses the sum of squared differences (SSD) as the criterion of patch matching, which can be written as

$$\psi_{\widehat{q}} = \arg\min_{\psi_q\in\Phi} d\left(\psi_{\widehat{p}}, \psi_q\right) \tag{4}$$

with

$$\begin{aligned} d\left(\psi_{\widehat{p}}, \psi_q\right) &= d_{SSD}\left(\psi_{\widehat{p}}, \psi_q\right) \\ &= \sum_{t\in\psi_{\widehat{p}}} \left[\psi_{\widehat{p}}(t)M_{\widehat{p}}(t) - \psi_q(t)M_{\widehat{p}}(t)\right]^2, \end{aligned} \tag{5}$$

where $M_{\widehat{p}}$ is a mask with the same size as $\psi_{\widehat{p}}$ and having value 1 if the point in $\psi_{\widehat{p}}$ is known and 0 otherwise. Thus only the valid known points are used for the patch matching.

After the most similar patch $\psi_{\widehat{q}}$ is found, the unknown points in $\psi_{\widehat{p}}$ are filled with the points in $\psi_{\widehat{q}}$ at the corresponding locations, as shown in Fig. 1(d). The confidence values also need to be initialized for the newly filled pixels, as

$$C(t) = C\left(\widehat{p}\right), \forall t \in \psi_{\widehat{p}} \cap \Omega. \tag{6}$$

After filling $\psi_{\widehat{p}}$, the boundary $\partial\Omega$ is renewed. The whole process is repeated until whole target region is filled.

3. Proposed approach

Criminisi method is effective in processing both structure and texture images but it still has limitations. A space varying updating strategy is proposed in Section 3.1.1 to decrease the dropping