Pattern Recognition Letters 32 (2011) 1256–1266

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Image inpainting with salient structure completion and texture propagation

Shutao Li *, Ming Zhao

College of Electrical and Information Engineering, Hunan University, Changsha 410082, China

ARTICLE INFO

Article history:
Received 10 May 2010
Available online 22 March 2011
Communicated by G. Sanniti di Baja

Keywords:
Image inpainting
Texture synthesis
Structure completion
Curve fitting
Curve extension

ABSTRACT

Image inpainting technique uses structural and textural information to repair or fill missing regions of a picture. Inspired by human visual characteristics, we introduce a new image inpainting approach which includes salient structure completion and texture propagation. In the salient structure completion step, incomplete salient structures are detected using wavelet transform, and completion order is determined through color texture and curvature features around the incomplete salient structures. Afterwards, curve fitting and extension are used to complete the incomplete salient structures. In the texture propagation step, the proposed approach first synthesizes texture information of completed salient structures. Then, the texture information is propagated into the remaining missing regions. A number of examples on real and synthetic images demonstrate the effectiveness of our algorithm in removing occluding objects. Our results compare favorably to those obtained by existing greedy inpainting techniques.

1. Introduction

Image inpainting, also known as image completion, aims to fill missing pixels in unknown regions of an image in visually plausible way (Bertalmio et al., 2000). Image inpainting is an important topic in image processing, which can be applied in many areas, such as computer graphics, image editing, film postproduction, image restoration. Although image inpainting technique is very useful, the inpainting task is far from being a trivial accomplishment. It is still a challenging problem in computer graphics and computer vision.

Image inpainting has attracted a considerable amount of researches in recent years. Roughly speaking, existing image inpainting approaches can be divided into three categories. The first category is partial differential equation (PDE) based method introduced by Bertalmio et al. (2000) and Chan and Shen (2001, 2002). These methods attempt to fill the missing regions of an image through a diffusion process which smoothly propagates information from boundary toward interior of the missing region. This diffusion process is simulated through solving a high order PDE. Bertalmio's method fills a hole in an image by propagating image information in the isophotes direction. Chan and Shen propose the total variational (TV) inpainting model inspired by Bertalmio's work. The TV inpainting model applies an Euler–Lagrange equation inside the inpainting region, and with simply employs anisotropic diffusion based on the contrast of the isophotes (Chan and Shen, 2001). The curvature-driven diffusion (CDD) model (Chan and Shen, 2002) extends the TV algorithm by considering geometric information of isophotes when defining the "strength" of the diffusion process, thus allowing the CDD method to proceed over larger areas. Dobrosotskaya and Bertozzi combine wavelets with variational methods to eliminate blurry edges (Dobrosotskaya and Bertozzi, 2008). These PDE based image inpainting techniques work at the pixel level, and perform well for small gaps. However, for larger missing regions or textured regions, the PDE based methods cannot maintain texture detail, which often leads to blurry artifacts. An inpainting example with Bertalmio's method is given in Fig. 1(b).

The second category is exemplar-based method, which tries to fill the missing regions by copying content from the existing part of the image. The exemplar-based method starting with the seminal work of Efros and Leung (1999) can be divided into two kinds, pixel level method (Wei and Levoy, 2000) and patch level method (Criminisi et al., 2004; Cheng et al., 2005; Ignácio and Jung, 2007). The difference is whether the textures are synthesized by one pixel or one patch at each step. Up to now, in the most famous exemplar-based inpainting method proposed by Criminisi et al. (2004), the inpainting order is determined by the structure information of the image. Their insight is that the source region should be propagated along linear isophotes first, thus making the inpainting reasonable regardless of the boundary shape of the inpainting region. Venkatesh et al. (2009) extend this approach to the video inpainting field. However, Criminisi's method is limited to inpaint linear structures, often resulting in discontinuities of curving structures. There are several improved exemplar-based inpainting approaches based on the Criminisi's method (Cheng et al., 2005; Ignácio and Jung, 2007). However, a major drawback of these exemplar-based approaches stems from their greedy way of filling the image, which can often lead to visual inconsistencies. To overcome this deficiency, Sun et al. synthesizes image patches along user-specified

* Corresponding author. Tel.: +86 731 88822924; fax: +86 731 88822224.
E-mail addresses: shutao_li@yahoo.com.cn (S. Li), mingzhao_frank@yahoo.cn (M. Zhao).

0167-8655/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2011.03.015

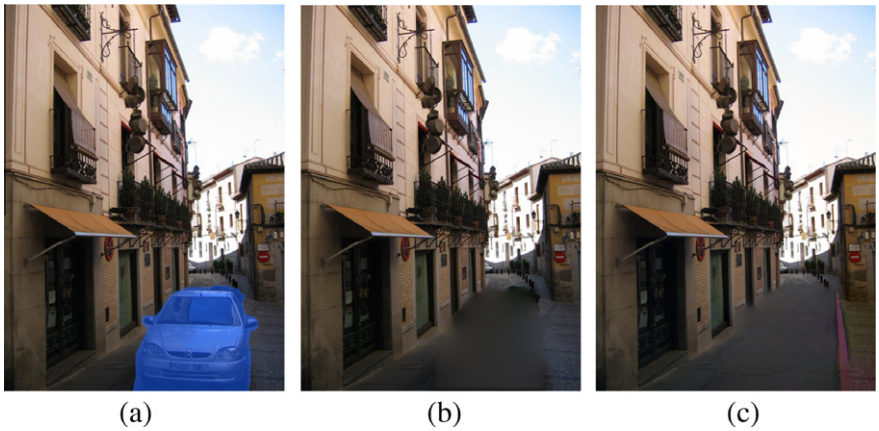

Fig. 1. An example of image inpainting: (a) target image; (b) inpainting result of Bertalmio's method (Bertalmio et al., 2000); and (c) inpainting result of Hays and Efros's method (Hays and Efros, 2008).

curves in the unknown region using patches selected around the curves in the source region (Sun et al., 2005). Although this method achieves excellent results, it requires user manually to specify important missing structure information by extending a few curves or line segments from the known to the unknown regions. Komodakis et al. propose an exemplar-based technique which considers the image completion problem as a discrete global optimization based on a Markov Random Field (Komodakis et al., 2007).

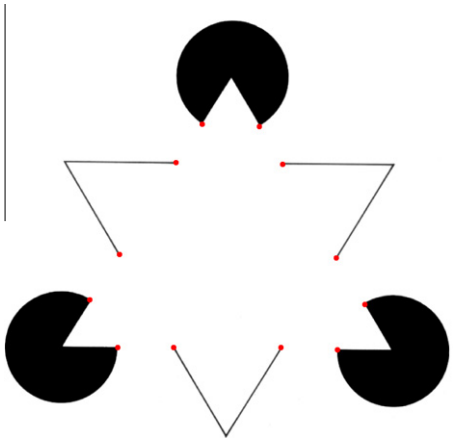

Fig. 2. Kanizsa triangle and T-junctions (marked with red point). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

However, Komodakis' method is time consuming. Different from the previous inpainting methods synthesizing structural and textural information from target image, Hays and Efros (2008) and Li et al. (2010), proposed novel image inpainting methods using a huge image database, which includes millions of photographs gathered from the Web. These image database based methods may alter the source region of target image. An example of Hays and Efros method is presented in Fig. 1(c).

The third category is combination method, which usually combines PDE with exemplar-based inpainting method together (Bertalmio et al., 2003; Grossauer, 2004). The combined method defines images that can be decomposed or segmented into structure component and texture component. Then the combination method treats the structure and texture components with different methods. Bertalmio et al. (2003) extend their PDE inpainting method (Bertalmio et al., 2000) to texture and structure according to the decomposition described above. After that the structure component is inpainted by their PDE inpainting method. Texture component is inpainted by Criminisi's method. Grossauer (2004) proposes a more extensive structural and textural inpainting framework, in which the image is firstly filtered using a curvature minimizing diffusion to generate structure component. Then, structure component is inpainted and segmented, and texture component is inpainted using the tree-based exemplar replication algorithm. These combination methods decompose or segment image into geometric features and texture. However, the image segmentation and decomposition is a challenge in image processing and the decomposition (segmentation) results are not often satisfactory. Moreover the combination method usually consists of PDE method

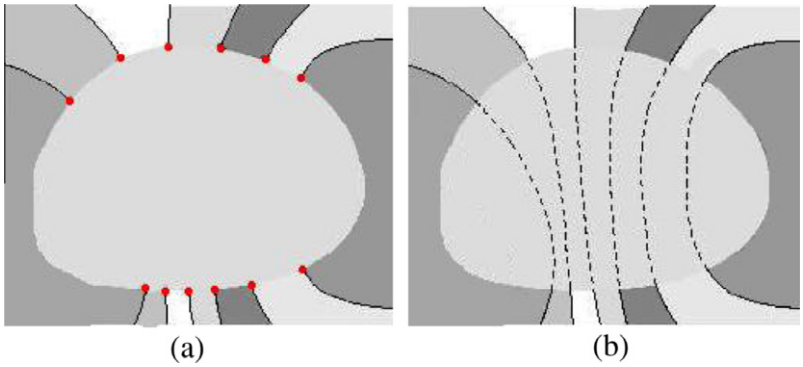

Fig. 3. A visual completion example: (a) an image with occlusion (T-junctions are marked with red point); (b) structures completion by visual prediction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and patch-based method. As we described above, the PDE method usually creates blurring effect.

Inspired by Sun's et al. (2005) and human vision mechanism in image occlusion detection and filling-in, we propose an inpainting method based on automatic incomplete salient structure detection and completion which improves the greedy inpainting method. The wavelet transform edge detection is utilized to detect incomplete salient structures. Then the incomplete salient structures are completed in the missing area through curve fitting and extension. The completed salient structures are used to guide exemplar-based texture propagation. Different from combination methods, the proposed method basically belongs to patch-based inpainting method. We use wavelet transform and curvature fitting to determine the inpainting order. Then the texture information is propagated by patch-based inpainting method. Compared to the combination method, our method does not need to decompositing the image into different parts and the results of our method are more robust. Different from greedy inpainting methods (Criminisi et al., 2004; Cheng et al., 2005; Ignácio and Jung, 2007), the proposed method completes the image salient structures first, thus avoiding the occurrence of visual inconsistencies during the image completion process. Also, the completed salient structures divide the source area and missing area into several regions. This not only reduces time consumption for texture synthesis, but also provides more precise texture information (Sun's et al., 2005).

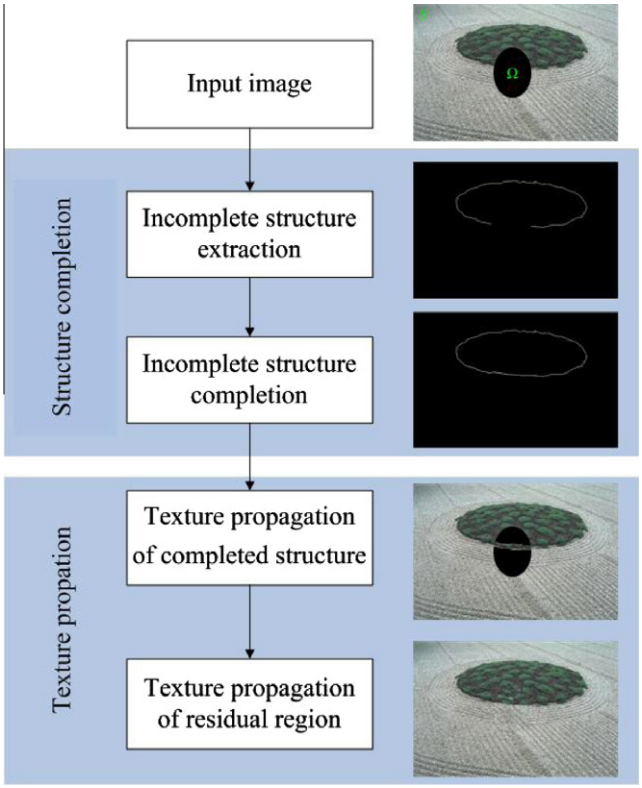

Fig. 4. Flowchart of the proposed method.

The rest of this paper is organized as follows. In Section 2, the characteristic of human vision in missing region detection and filling is briefly discussed. In Section 3, the detail of automatic incomplete salient structure detection and completion inpainting method is introduced. The experimental results are shown in Section 4. Finally, conclusions and discussions are provided in Section 5.

2. Inspiration from human vision characteristic

One of the challenges of image inpainting is how to fill missing area in a visually plausible way. However, it is a subjective criterion and difficult to be quantized. There is an example shown in Fig. 1. As can be seen, in visual criterion Fig. 1(c) is better

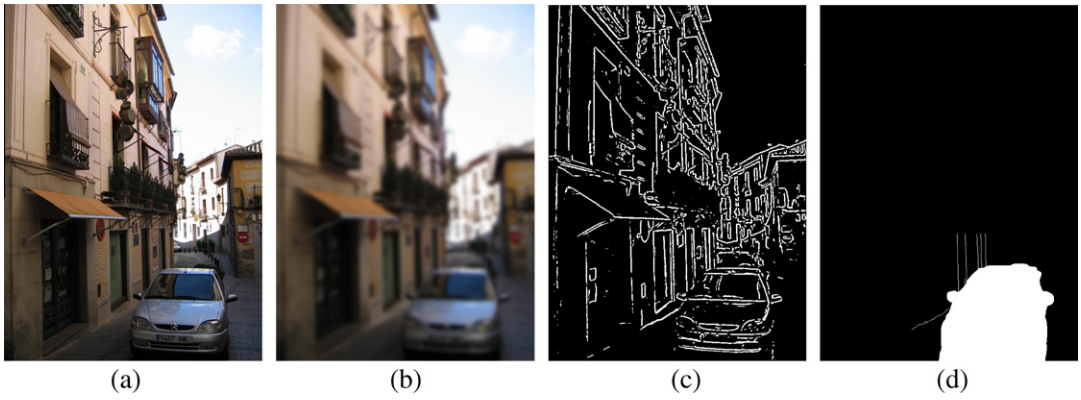

Fig. 5. An example of incomplete salient structure extraction: (a) target image; (b) blurred image of (a); (c) Salient structures of (a); and (d) incomplete salient structures. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

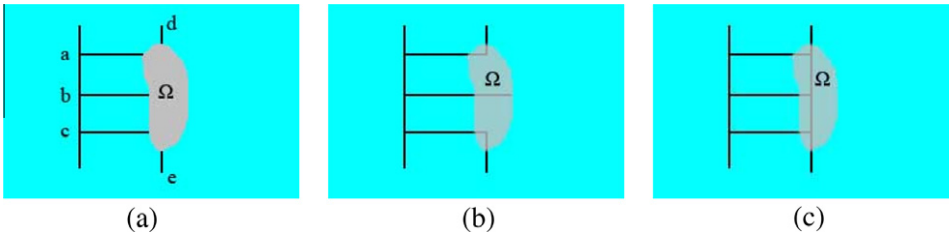

Fig. 6. An error example of structure completion: (a) a structure missing image; (b) a simple structure connection result; (c) the expected result.

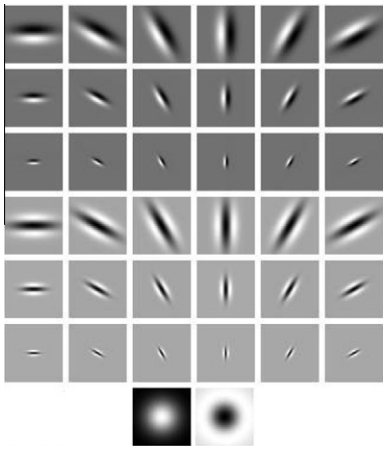

Fig. 7. The maximum response 8 filters.

Fig. 8. An example of incomplete salient structures completion: (a) the incomplete salient structures; (b) the completed missing structures by our approach.

than Fig. 1(b). But the mean square error between inpainted image shown in Fig. 1(b) and the original image is smaller than Fig. 1(c) ($MSE_b = 50.1$; $MSE_c = 52.3$). Therefore, we should consider the nature of human vision in image missing detection and filling.

From a series of perceptual experiments, Kanizsa found that human vision system perceives image missing areas through detecting T-junctions (Kanizsa, 1985; Pessoa et al., 1998). The T-junctions are points where structure of one object abruptly intersects with that of another object (or the boundary of missing region) and forms a junction in the shape of the letter "T". Once the T-junctions are perceived, human vision system detects the incomplete structures simultaneously. Then, human brain performs a continuation of incomplete structure between T-junctions. This can be proven by Fig. 2, the famous Kanizsa triangle. When the image is complex, human visual system will find similar incomplete structures as a pair and believe they should be connected as one (Fig. 3).

Moreover, Nill and Bouzas pointed out that human vision is liable to ignore large uniform area and sensitive to structural region (Nill and Bouzas, 1992). For example, although the street area of the Fig. 1(b) and (c) are blurred, the completed building structure and the clear street boundary lead to a visual advantage that makes us to believe that Fig. 1(c) is better than Fig. 1(b). Therefore, to achieve an outstanding inpainting result, the key is to find defective structures and complete them.

3. Salient structure completion algorithm

The inpainting methodology of the exemplar-based method can be defined as follows: given an incomplete image I, as well as the target (missing) region Ω, and source region S where S is a subset of I/Ω, the inpainting process is to fill the target region Ω by propagating information from S. Fig. 4 provides an overview of the proposed inpainting method, which is explained in details as follows.

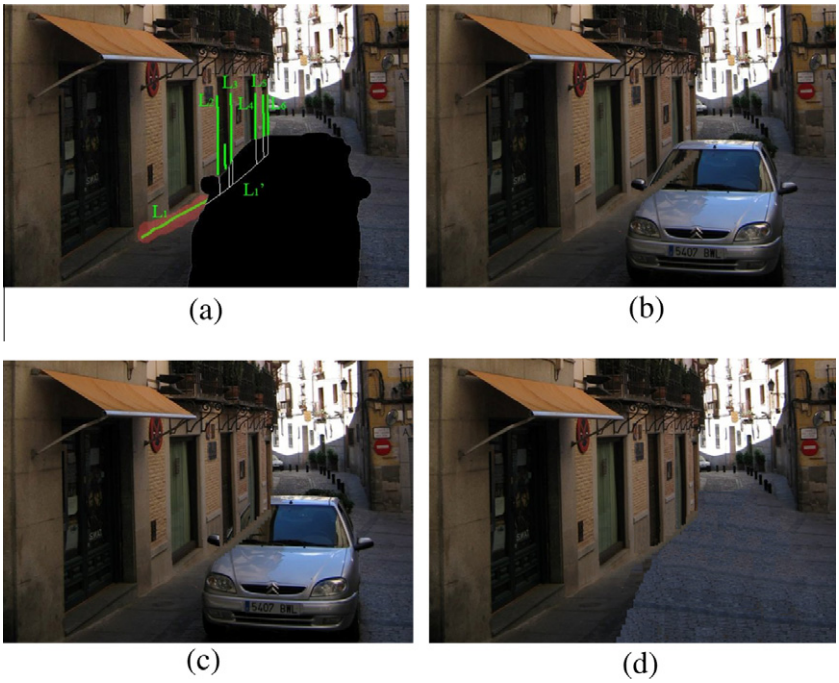

Fig. 9. The texture propagation results: (a) target image; (b) texture propagation of L_1'; (c) result of completed missing structures texture propagation; and (d) the final result.

3.1. Incomplete salient structure completion

3.1.1. Incomplete salient structure extraction

In our opinion, the structural integrity of the image is the basis to obtain excellent inpainting results. As inpainting images usually contain large number of structures while the salient structure is limited. Therefore, our approach pays attention to the integrity of salient structures. The salient structures are generally the most important edges of the image. Different from other edges, the salient structures allow us to immediately concentrate on objects of interest in the image. However, the salient structures are often confused with a cluttered background. In order to extract the incomplete salient structures, texture information and less significant edges should be ignored. Therefore, we firstly implement a Gaussian blur operation on the original image and the kernel is given in following equation:

$$G(x,y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}. \quad (1)$$

An example of Gaussian blur operation is shown in Fig. 5(b), where the radius of Gaussian operation is 5 and standard deviation σ is 2.

Then, we use wavelet transform to extract the salient structures. In images, salient structures are usually created by different local intensity profile such as occlusions, shadows, textures. Therefore, in order to label the salient structure precisely, it is necessary to analyze the singularities of the local properties. In mathematics, singularities are generally characterized by their Lipschitz exponents. According to the wavelet theory, the Lipschitz exponents can be computed from the evolution across scales of the wavelet transform local extrema, which correspond to zero crossings of its derivative (Mallat and Zhong, 1992).

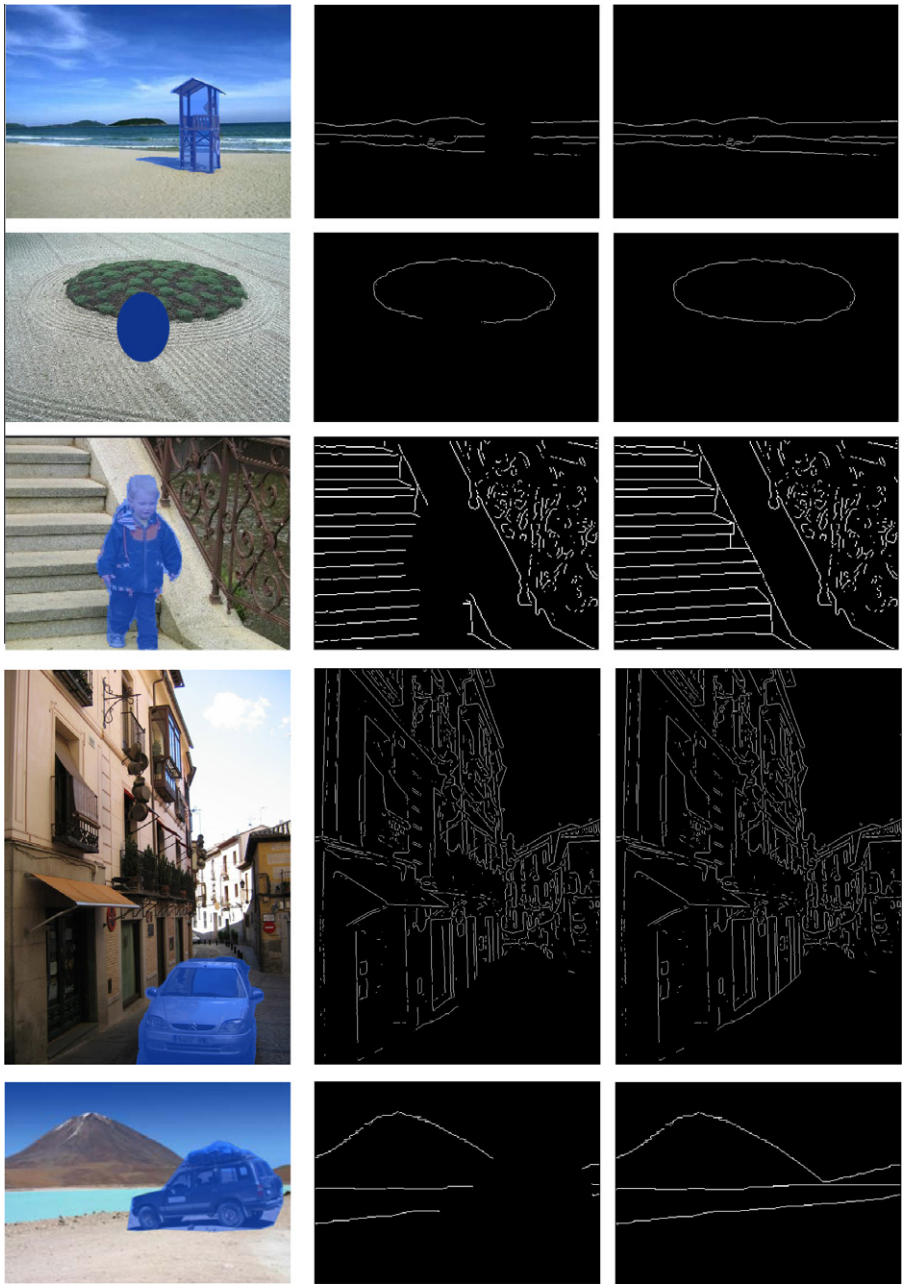

Fig. 10. Some results of salient structure completion.

The wavelet edge detectors smooth the signal at various scales and detect sharp variation points from their first-order or second-order derivatives. The extrema of the first-order derivative corresponds to the zero crossings of the second-order derivative. The function $\theta(x, y)$ is quadratic differentiable and we defines, $\psi^a(x, y)$ and $\psi^b(x, y)$ as the first-order and second-order derivative of $\theta(x, y)$.The definition of $\psi^a(x, y)$, $\psi^b(x, y)$ and $\theta(x, y)$ can be found in (Mallat and Zhong, 1992). A salient structure extraction result of Fig. 5(a) is shown in Fig. 5(c).

The incomplete salient structures are the detected edges which intersect with missing regions and we store the coordinates of the incomplete salient structures for the following curve features calculation. An example of extracted incomplete salient structures map is shown in Fig. 5(d).

3.1.2. Incomplete salient structure completion

Connecting incomplete salient structures is the key to obtain creditable inpainting results. However, simple connection or extension of the incomplete salient structures usually leads to negative results. For example, in Fig. 6(b) the restored salient missing structure b passes through the missing region that makes the inpainting result unsatisfactory. Therefore, we should connect the incomplete salient structures to human visual characteristic more approximately. As we discussed in Section 2, the human visual system completes the incomplete structures by considering similarity their similarity. According to this characteristic, we propose two rules in the incomplete structure completion algorithm:

(a) Completeness: All the extended incomplete salient structures should intersect at the target area to ensure that none of the restored structures passes through the target area.
(b) Continuation order: The similar incomplete salient structures are marked up as a pair and connected at first. The unpaired incomplete salient structures should be extended into the target region until they intersect with the completed paired structures.

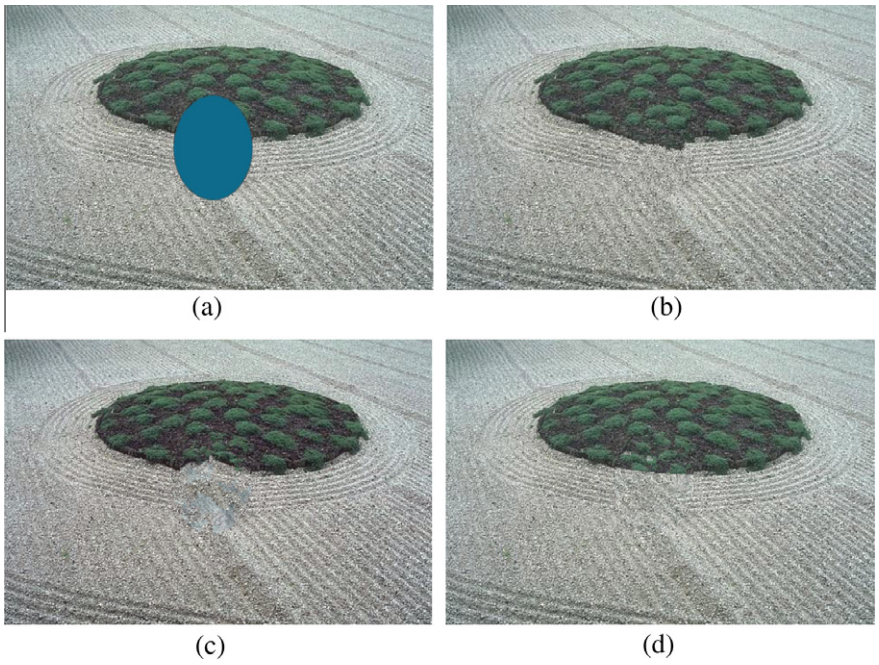

Fig. 11. Comparison of missing information filling, patch size $n = 15$: (a) target image; (b) result of the Ignácio and Jung's method (Ignácio and Jung, 2007); (c) result of the Criminisi's method (Criminisi et al., 2004); and (d) result of the proposed method.

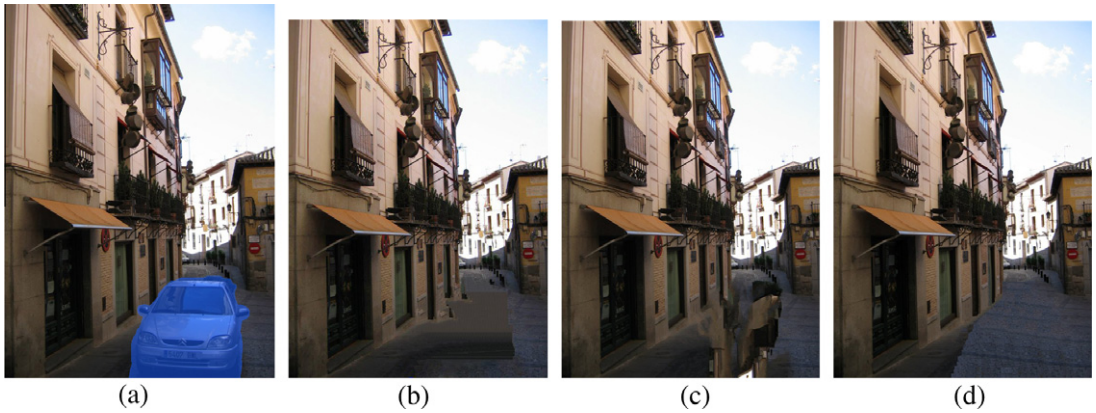

Fig. 12. Removal of the car, patch size $n = 19$: (a) original image; (b) result of the Ignácio and Jung's method (Ignácio and Jung, 2007); (c) result of the Criminisi's method (Criminisi et al., 2004); and (d) result of the proposed method.

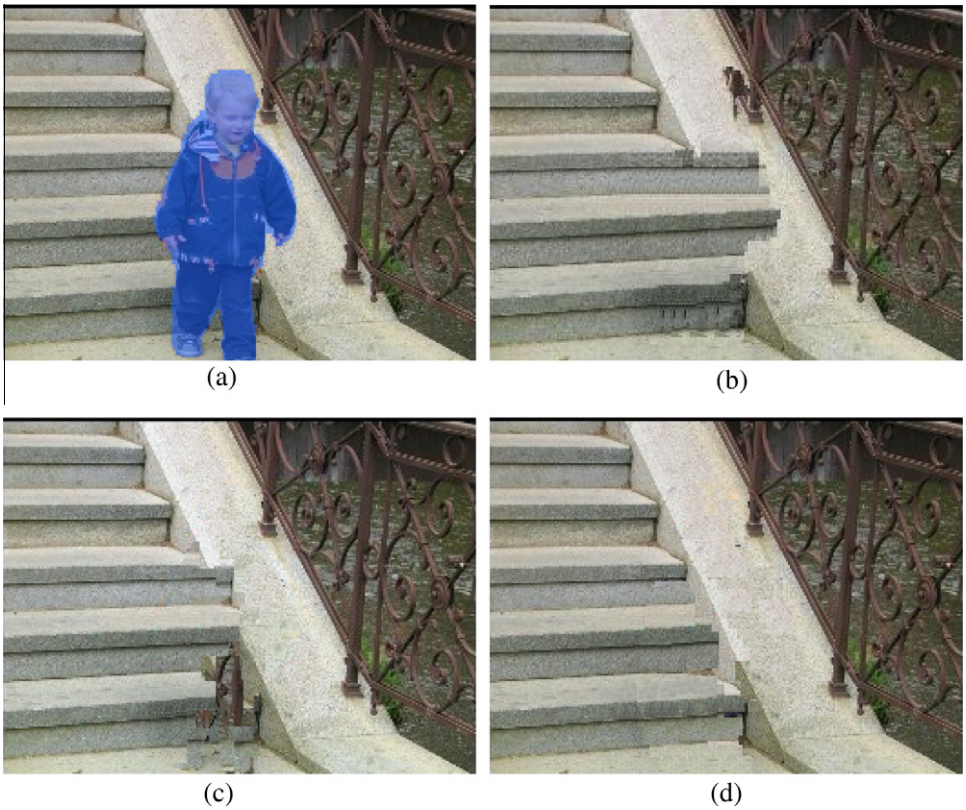

Fig. 13. Removal of the child, patch size $n = 13$: (a) original image; (b) result of the Ignácio and Jung's method (Ignácio and Jung, 2007); (c) result of the Criminisi's method (Criminisi et al., 2004); and (d) result of the proposed method.

Inspired by content-based image retrieval (CBIR), color, texture and curvature features are used to determine the similarity of incomplete salient structures. Color feature describes the local color composition in terms of spatial domain around the incomplete structures. Texture feature describes the spatial characteristics of the grayscale component around the incomplete structures and curvature feature is used to describe the geometrical similarity of the salient structures. They provide a simple but effective characterization of incomplete salient structures, which are used to pair the incomplete salient structures.

The color and texture features of the incomplete salient structures are calculated from a surrounding area of the source region. The surrounding region size indicates how many pixels should be used to measure the texture and color similarity. In our experiments, we find that the surrounding region with 50×50 pixels is appropriate. To measure color distance, we choose the improved optimal color composition distance method (IOCCD) (Chen et al., 2005). The IOCCD is based on an important characteristic of human color perception, i.e., the human eye cannot simultaneously perceive a large number of colors. Thus, the IOCCD method is effective and consistent with human visual characteristic. The color distance D_1 is calculated as follows:

$$D_1 = IOCCD(C_{R1}, C_{R2}), \tag{2}$$

where the color feature vectors C_{R1} and C_{R2} imply the RGB color composition of two surrounding regions.

As for the texture similarity, the MR8 (maximum response) filters are used (Varma and Zisserman, 2009). The MR8 filter consists of 38 filters and only 8 filter responses as shown in Fig. 7. The filters include a Gaussian filter, a Laplacian of Gaussian (LOG) filter, an edge (first derivative) filter and a bar (second derivative) filter which both have six orientations and three scales$(\sigma_x, \sigma_y) = \{(1, 3), (2, 6), (4, 12)\}$. The responses of the isotropic filters (Gaussian and LOG) are used directly. The responses of the oriented filters (bar and edge) are "collapsed" at each scale by only using the maximum filter responses across all orientations. This gives eight filter responses in total and ensures that the filter responses are rotationally invariant. The texture distance D_2 is calculated as follows:

$$D_2 = \frac{|MR8(T_{R1}) - MR8(T_{R2})|_F}{250}, \tag{3}$$

where the T_{R1} and T_{R2} imply the intensity of two surrounding regions.

In addition to color and texture features, we also consider the curvature of the incomplete salient structures. In the incomplete salient structures detection step, we extract all the incomplete salient structures and store every pixel of these salient structures. Therefore, we can easily determine whether these incomplete salient structures are straight lines with standard Hough transform. After that the slope of those straight incomplete salient structures are calculated to determine whether they are parallel. Then the

Table 1
Consuming time comparison of different inpainting method(s).

Method	Fig. 11 (480×320 pixels)	Fig. 12 (580×800 pixels)	Fig.13 (260×200 pixels)
Ignácio and Jung's method (Ignácio and Jung, 2007)	212.2	825.1	101.7
Criminisi's method (Criminisi et al., 2004)	221.5	864.3	103.9
The proposed method	178.5	647.4	97.6

incomplete salient structures pairing problem can be defined as follows:

$$D(i,j) \triangleq IP \cdot \left[\alpha \cdot D_1(\mathbf{f}_C^i, \mathbf{f}_C^j) + \beta \cdot D_2(\mathbf{f}_T^i, \mathbf{f}_T^j) + \gamma \cdot D_3(\mathbf{f}_{Cur}^i, \mathbf{f}_{Cur}^j)\right], \quad (4)$$

$$Paired(i,j) = \arg \min_{\substack{i \neq j \ IP=1 \\ i,j \in \{1,\ldots,N\}}} D(i,j), \quad (5)$$

$$D_3(\mathbf{f}_{Cur}^i, \mathbf{f}_{Cur}^j) = |curvature_i - curvature_j|, \quad (6)$$

where i and j are indexes of two different incomplete salient structures. N is the number of incomplete salient structures. IP is a logical variable, if i parallel to j, IP is equal to 1; else IP is equal to 0. D_1, D_2 and D_3 are the IOCCD, MR8 and curvature distance of i and j. α, β and γ are used to balance color distance, texture similarity and curvature difference of i and j. In our method, the color and texture features are superior to curvature feature. In most instances, color similarity plays a decisive role. Therefore, we set $\alpha > \beta > \gamma$ in our experiments. Considering the number of N is small, we solve the above optimization problem by checking all the possible combinations.

After the incomplete salient structures pairing procedure, the paired incomplete salient structures are completed based on the curve equations which are calculated through second-order polynomial curve fitting. We firstly estimate the curve equation of every incomplete salient structure through second-order polynomial fitting. Then the paired incomplete salient structures are extended into the missing region until they intersect inward the missing region according to the curve equation. Note that if the extension parts of the paired incomplete salient structures do not intersect in the missing region then we connect them in the nearest position. Afterwards, we extend the unpaired incomplete salient structures until they intersect with the restored part of the paired incomplete salient structures. An example of incomplete salient structure completion is shown in Fig. 8.

3.2. Texture propagation

3.2.1. Texture propagation of completed salient structures

After structure completion, we propagate texture information into the target region through Criminisi's patch based inpainting method (the patch size is $n \times n$ pixels). The Criminisi's algorithm fills the texture information by computing a prior target patch. The prior target patch $p(t)$ is defined as follows:

$$p(t) = C(p) \times D(p), \quad (7)$$

$$C(p) = \frac{\sum_{q \in \psi_p \cap S} r(p)}{size(\psi_p)}, \quad (8)$$

$$D(p) = |\langle \nabla^{\perp} I(p), n(p) \rangle|, \quad (9)$$

where $D(p)$ is data term and $C(p)$ is confidence term. Ψ_p is the patch with p to be the center pixel. $size(\Psi_p)$ is the number of pixels in patch Ψ_p. $r(p)$ is assigned to each pixel $p(x,y)$ to indicate its reliability. During initialization, $r(p)$ is set to 1 in source region and 0 in the target region. $n(p)$ is the normal of Ω boundary, $\nabla I(p)$ is the isophotes which are perpendicular to gradient direction $\nabla I(p)$.

However, applying texture synthesis algorithm directly may generate poor results. The texture synthesis algorithm may produce irrelevant texture information from the entire source region.

Fig. 14. Comparison on the examples in Sun's method (Sun et al., 2005): (a) original images; (b) unknown regions and human provided structures; (c) completion results from (Sun et al., 2005); and (d) our results.

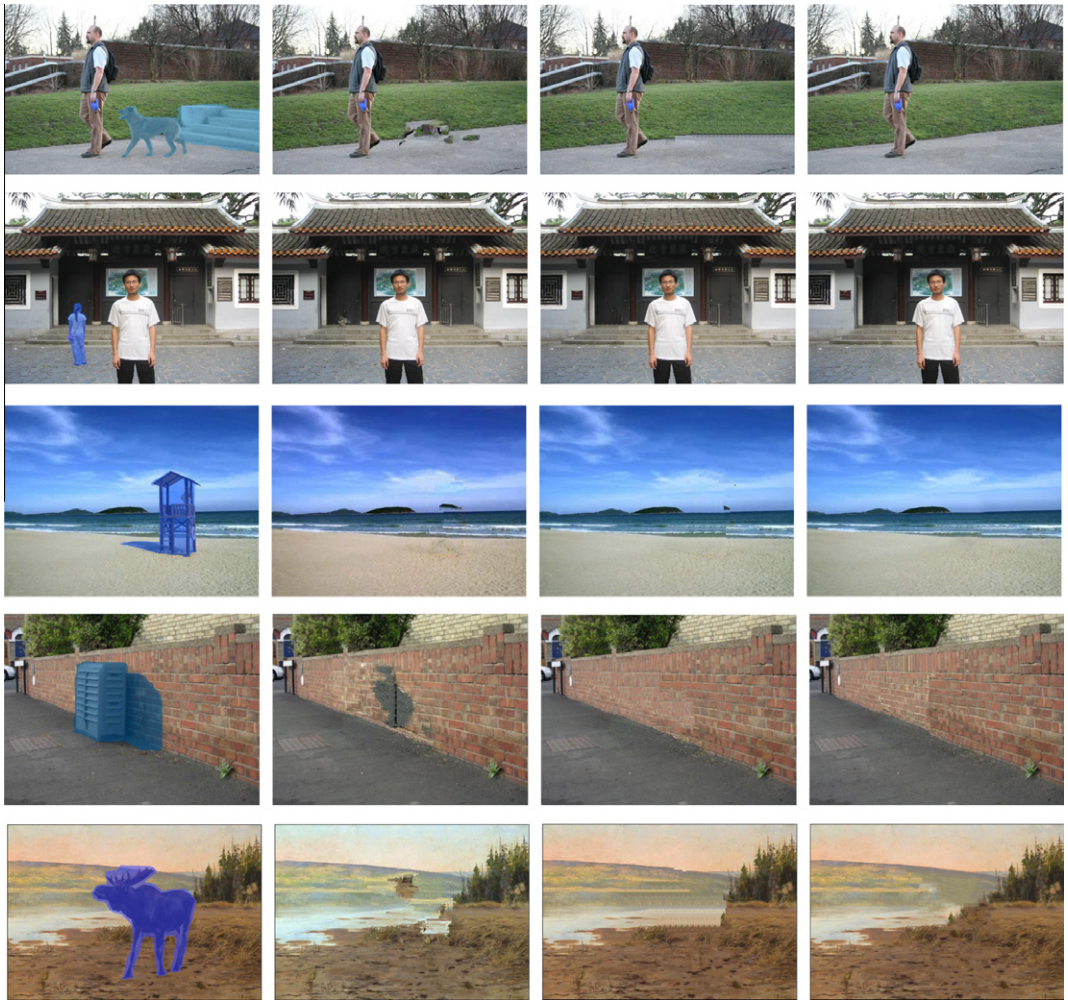

Fig. 15. Additional results (the target regions are marked in blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To overcome this deficiency, we first synthesize the texture information of the completed missing structures (white lines in Fig. 9(a)). The texture information of the completed missing structures should be copied from the corresponding incomplete structures. For example, in Fig. 9(a) the corresponding incomplete structure of completed missing structure L_1' is L_1. For each completed structure, the source region should be pruned. To prune source region for each completed missing structure, a morphological dilation operation is used on the corresponding incomplete structure. The structuring element of morphological dilation is a $2n \times 2n$ matrix of all '1'. The pruned source region is the intersection part of the original source region and the dilation operation result (the pruned source region of L_1' is shown in Fig. 9(a) marked with brown). The Criminisi's method successfully propagates texture information for completed salient structures by pruning the source region. The result of texture information propagation of L_1' is shown in Fig. 9(b). The texture of completed salient structure L_1' has been propagated from L_1 successfully. After that the textures of remnant completed salient structures are propagated in the same way. The result of texture propagation of other incomplete salient structures (L_2, L_3, L_4, L_5, L_6) is shown in Fig. 9(c).

3.2.2. Texture propagation of residual region

After the texture information of completed missing structures has been propagated, there are still large target regions that need to be filled. Unlike greedy patch-by-patch methods (Criminisi et al., 2004; Ignácio and Jung, 2007), the proposed method fills the residual texture information in a more effective way. As described in (Sun et al., 2005), the completed salient structures (or their extension) divide the target region and source region into several sub-regions. Each target sub-region is usually adjacent to one source sub-region. Therefore, we can synthesize the residual texture information for each target sub-region from the corresponding adjacent source sub-region. If there is more than one adjacent source sub-region, they should be merged as one sub-region. With Criminisi's method, texture information of target sub-region can be reliably and efficiently propagated from the corresponding adjacent sub-regions. As an example, the final inpainting result of Fig. 9(a) is presented in Fig. 9(d).

4. Experimental results

In this section, inpainting experiments are performed in the environment of an AMD Sempron 2600+ CPU 1.8GHz with a 512 MB RAM PC, operated under Windows XP and Matlab 7.0. The proposed method has four parameters: size of the texture synthesis patch n and weight coefficients α, β and γ in Eq. (6). Criminisi et al. (2004) proposed that the patch size n should be set slightly larger than the largest distinguishable texture element, or "texel," of the source region. Therefore, in our experiments, we manually

set the patch size n to be greater than the largest texel of the source region.

4.1. Experiments of salient structures completion

One advantage of the proposed inpainting method is that we complete the salient structures by considering the color, texture and curvature features of them. The parameters α, β and γ are used to balance these three features. As stated in Section 3.1.2, the color is superior to texture and the curvature feature, it works when the difference of color and texture features is very small. In the following experiments, the parameters α, β and γ in Eq. (13) are always set to 0.52, 0.38 and 0.1. However, we must note that β should be larger than α when the color of the target image is homochromy. For example, in the snow photos white is the dominating color. In this case, the texture parameter β is more important than color parameter α.

Some original images are listed in the first column of Fig. 10 and the target areas are marked in blue. The corresponding salient structure images extracted by our approach are presented in the second column. These salient structures completed through the proposed method and the corresponding results are shown in the third column of Fig. 10. As can be seen, the proposed salient structure completion approach works well.

4.2. Experiments of image inpainting

We compare the proposed inpainting method with two representative methods. One is Ignácio and Jung's method (Ignácio and Jung, 2007), i.e., a one level decomposition wavelet based patch-by-patch image inpainting algorithm, with minimum patch size of 8×8 pixels and maximum patch of 16×16 pixels. The other is the Criminisi's exemplar-based method (Criminisi et al., 2004), which is the most well-known exemplar-based inpainting method and we manually set the patch size greater than the largest structure in the image.

We first compare these inpainting methods on an artificial occlusion image shown in Fig. 11(a), which contains a hole in the intersection part of the image. We can see from Fig. 11(b), the Ignácio's method completes the salient structure correctly. But this method produces an unexpected protuberance. The Criminisi's exemplar-based method brings inaccurate structure information, because the curvature of incomplete salient structure is large. It can be seen Fig. 11(d) that our inpainting method not only restores more structural information than Fig. 11(b) and (c), but also reconstructs the texture as well.

In the following experiments, we want to remove objects from real scene pictures. Each of them contains rich texture and structure information. Figs. 12(b) and 13(b) are the inpainting results obtained by the Ignácio and Jung's method. In Fig. 12(b), the image is blurred with the Ignácio and Jung's inpainting method. In Fig. 13(b), the Ignácio and Jung's method produces inaccurate structures. The Criminisi's method can fill the target region without blurring (Figs. 12(c) and 13(c)). However, some undesirable objects are generated during the inpainting process, which lead to implausible inpainting results.

Our inpainting results, shown in Figs. 12(d) and 13(d), are more reasonable and natural than those of the Ignácio and Jung's method or Criminisi's method. These inpainting results show that the proposed algorithm is the best one among them.

We provide a comparison of speed on Figs. 11–13. As shown in Table 1, the proposed method is less time-consuming when compared with other two methods. And the consuming time required is related to the size of the input image and the area of missing region.

In Fig. 14, we compare our algorithm with Sun's semi-automatic inpainting algorithm (Sun et al., 2005). The value of n is 11, 11 and

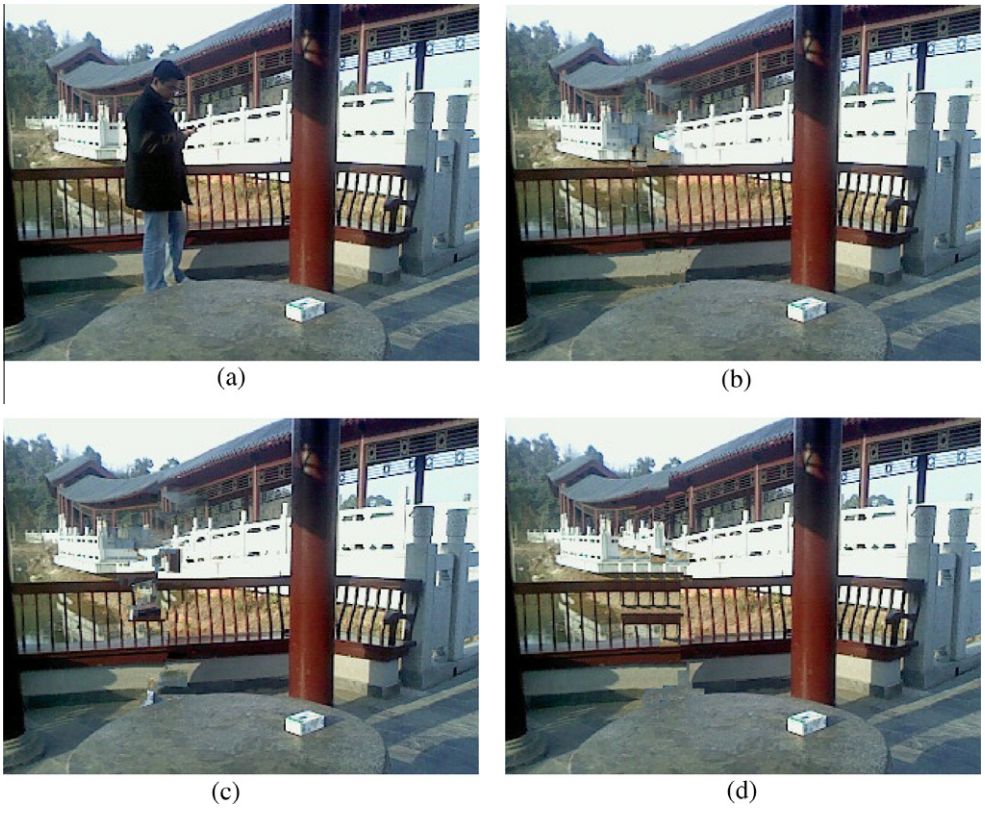

Fig. 16. A failing example of the proposed method, patch size $n = 11$: (a) original image; (b) result of the proposed method; (c) result of the Criminisi's method (Criminisi et al., 2004); (d) result of the Ignácio and Jung's method (Ignácio and Jung, 2007).

13 from the first row to the third row of Fig. 14, respectively. The Sun's inpainting algorithm resolves the structures in the missing region utilizing the human-labeled structures as the guidance, and achieves state-of-the-art results for completing the complex structures. Whereas, our method automatically completes the missing structures and synthesizes texture reasonably without any user intervention. As shown in Fig. 14, the inpainting results of our method are comparable to the results in (Sun et al., 2005), which uses human-labeled structures as guidance.

Moreover, Fig. 15 presents further results on image inpainting. The first column of Fig. 15 is original images and the inpainting targets are marked as blue. The second and third columns are the inpainting results of Ignácio and Jung's (Ignácio and Jung, 2007) method and Criminisi's method (Criminisi et al., 2004). The value of n is 15, 19, 13, 15, and 13 from the first row to the fifth row of Fig. 15, respectively. The results of our method are presented in the fourth column of Fig. 15 and these results along with those in Figs. 11–14 demonstrate the effectiveness of the proposed method. As can be seen from the presented examples, our approach successfully completes smooth areas, textured areas and structured areas. At this point, our method can regenerate texture and structure information successfully. The inpainting results are apparently consistent with the human visual perception.

The above experiments show that the proposed method is effective for the images with prominent structures. However, when it comes to the over-complex structures (edges), our method will leads to undesired results. A failing example is given in Fig. 16. As can be seen from this figure, our approach works well on the restoration of pavilion but leads to visual incoherent in the cloister, because the structures of cloister are too complex and it is difficult to extract and complete salient structure of the cloister. But the proposed method is still better than other two methods, especially in the pavilion part.

5. Conclusions and discussions

In this paper, we have presented a novel inpainting method based on automatic salient structure completion. Using incomplete salient structure extraction and completion, our method completes the structure of missing region reasonably. The completed salient structures divide the target area into several sub-regions. Then, texture propagation is used to synthesize the texture information with samples from corresponding adjacent sub-regions. This reduces the running-time and offers more precise texture information. The proposed method does not need any user intervention and provides similar good results as Sun's method (Sun et al., 2005). Moreover, the proposed method is equivalent to Criminisi's method when there is no incomplete salient structure in the target image.

Although the proposed method performs well in most circumstances, one limitation in our approach is that the incomplete structure completion approach produces unreasonable structures when the structure of the input image is too complex. Therefore, in the future, we will study how to build a comprehensive model to deal with this deficiency.

Acknowledgments

The authors would like to thank the editor and anonymous reviewers for their detailed review, valuable comments, and constructive suggestions. This paper is supported by the National Natural Science Foundation of China (Nos. 60871096 and 60835004), the Ph.D. Programs Foundation of Ministry of Education of China (No. 200805320006), the Key Project of Chinese Ministry of Education (2009-120), and the Open Projects Program of National Laboratory of Pattern Recognition, China.

References

Bertalmio, M., Sapiro, G., Caselles, V., et al., 2000. Image inpainting. In: Proc. Computer Graphics (SIGGRAPH'00), Singapore, pp. 417–424.
Bertalmio, M., Vesa, L., Sapiro, G., et al., 2003. Simultaneous structure and texture image inpainting. IEEE Trans. Image Process. 12 (8), 882–889.
Chan, T.F., Shen, J., 2001. Non-texture inpainting by curvature driven diffusion. J. Vision Commun. Image Represent. 12 (4), 436–449.
Chan, T.F., Shen, J., 2002. Mathematical models for local non-texture inpaintings. SIAM J. Appl. Math. 62 (3), 1019–1043.
Chen, J.Q., Pappas, T.N., Mojsilovic, A., et al., 2005. Adaptive perceptual color-texture image segmentation. IEEE Trans. Image Process. 14 (10), 1524–1536.
Cheng, W. H., Hsieh, C.W., Lin, S.K., et al., 2005. Robust algorithm for exemplar-based image inpainting. In: Proc. Internat. Conf. on Computer Graphics, Imaging Vision, Beijing, pp. 64–69.
Criminisi, A., Perez, P., Toyama, K., 2004. Region filling and object removal by exemplar-based inpainting. IEEE Trans. Image Process. 13 (9), 1200–1212.
Dobrosotskaya, J.A., Bertozzi, A.L., 2008. A wavelet-laplace variational technique for image deconvolution and inpainting. IEEE Trans. Image Process. 17 (5), 657–663.
Efros, A.A., Leung, T.K., 1999. Texture synthesis by non-parametric sampling. In: Proc. Internat. Conf. on Computer Vision (ICCV'99), Kerkyra, pp. 1033–1038.
Grossauer, H., 2004. A combined PDE and texture synthesis approach to inpainting. In: Proc. European Conf. on Computer Vision, Slovansky ostrov, vol. 4, pp. 214–224.
Hays, J., Efros, A.A., 2008. Scene completion using millions of photographs. Commun. ACM 51 (10), 87–94.
Ignácio, U.A., Jung, C.R., 2007. Block-based image inpainting in the wavelet domain. Visual Comput. 23 (9–11), 733–741.
Kanizsa, G., 1985. Seeing and thinking. Acta Psychol. 59 (1), 23–33.
Komodakis, N., Tziritas, G., 2007. Image completion using efficient belief propagation via priority scheduling and dynamic pruning. IEEE Trans. Image Process. 16 (11), 2649–2661.
Li, H., Wang, S., Zhang, W., et al., 2010. Image inpainting based on scene transform and color transfer. Pattern Recognition Lett. 31 (7), 582–592.
Mallat, S.G., Zhong, S., 1992. Characterization of signals from multi-scale edges. IEEE Trans. Pattern Anal. Machine Intell. 11 (7), 710–732.
Nill, B., Bouzas, B., 1992. Objective image quality measure derived from digital image power spectra. Opt. Eng. 31 (44), 813–825.
Pessoa, L., Thompson, E., Noë, A., 1998. Find out about filling-in a guide to perceptual completion for visual science and the philosophy of perception. Behavioral Brain Sci. 21 (6), 723–802.
Sun, J., Yuan, L., Jia, J., et al., 2005. Image completion with structure propagation. ACM Trans. Graphics 24 (3), 861–868.
Varma, M., Zisserman, A., 2009. Statistical approach to material classification using image patch Exemplars. IEEE Trans. Pattern Anal. Machine Intell. 13 (11), 2032–2047.
Venkatesh, M.V., Cheung, S.C.S., Zhao, J., 2009. Efficient object-based video inpainting. Pattern Recognition Lett. 30 (2), 168–179.
Wei, L.Y., Levoy, M., 2000. Fast texture synthesis using tree-structured vector quantization. In: Proc. Computer Graphics (SIGGRAPH'00), New Orleans, pp. 479–488.