CrossMark

Striped-texture image segmentation with application to multimedia security

Jinxiong Ren[1] · Gang Chen[1] · Xiaoyan Li[2] · Kuang Mao[2]

Abstract In today's rapid growth of volume of multimedia data, security is important yet challenging problem in multimedia applications. Image, which covers the highest percentage of the multimedia data, it is very important for multimedia security. Image segmentation is utilized as a fundamental preprocessing of various multimedia applications such as surveillance for security by breaking a given image into multiple salient regions. In this paper, we present a new image segmentation approach based on frequency-domain filtering for images with stripe texture, and generalize it to lattice fence images. Our method significantly reduces the impact of stripes on segmentation performance. The approach proposed in this paper consists of three phases. Given the images, we weaken the effect of stripe texture by filtering in the frequency domain automatically. Then, structure-preserving image smoothing is employed to remove texture details and extract the main image structures. Last, we use an effective threshold method to produce segmentation results. Our method achieves very promising results for the test image dataset and could benefit a number of new multimedia applications such as public security.

✉ Jinxiong Ren
  renjx@zju.edu.cn

  Gang Chen
  cg@zju.edu.cn

  Xiaoyan Li
  kricellee@gmail.com

  Kuang Mao
  hzmaokuang@corp.netease.com

[1] College of Computer Science and Technology, Zhejiang University, CaoGuangBiao Building, 38 Zheda Road, Hangzhou 310027, People's Republic of China

[2] Hangzhou Research Institute, Netease(hangzhou), Inc., 599 Wangshang Road, Hangzhou 310052, People's Republic of China

✆ Springer

1 Introduction

With recent progress in computing technologies, the volume of high-resolution high-quality multimedia data such as images, video clips, animations, graphics, and audio have been growing exponentially over the past several years. Various multimedia modeling methods are proposed to process multimedia data in different area such as image categorization [25, 31, 32]. In [28], authors propose a feature selection algorithm to filter out the low efficiency features towards fast speech emotion recognition. Wang et al. propose a collaborative sparse coding framework that optimizes the classifiers and dictionary collaboratively for action recognition [20]. Zhang et al. develop a new semantically aware photo retargeting that shrinks a photo according to region semantics, with a mechanism transferring semantics of noisy image labels into different image regions [26]. Weakly supervised fixations prediction which leverages image labels to improve accuracy of human fixations prediction is proposed in [27], which can facilitate many multimedia applications, e.g., image retrieval, action recognition, and photo retargeting.

Image segmentation is a fundamental and widely studied problem which plays an important role in various multimedia applications, such as image annotation [15] and retrieval [6], object recognition [1] and matching [8], scene analysis [10], visual tracking [11] and social media mining [18], and security screening [7]. It is also very important due to its great benefit in multimedia security.

Under different motivations, various image segmentation algorithms have been developed, which have achieved very promising performance. Vese and Chan proposed a new multiphase level set framework for image segmentation using the Mumford and Shah model, for piecewise constant and piecewise smooth optimal approximations [19]. Comaniciu and Meer's Mean Shift seeks the modes of a non-parametric probability distribution in a feature space, and appears to well respect object details though it tends to split an object into pieces [5]. Shi and Malik's Normalized Cuts (Ncut) treats image segmentation as a graph partitioning problem and for segmenting the graph, propose the normalized cut criterion which measures both the total dissimilarity between the different groups as well as the total similarity within the groups [17]. Boykov and Jolly proposed an algorithm for general purpose interactive segmentation of N-dimensional images [3]. In [30], graphlets are introduced to represent a photo's aesthetic features, a probabilistic model is proposed to transfer aesthetic features from the training photo onto the cropped photo. Zhang et al. present a weakly supervised image segmentation algorithm by learning the distribution of spatially structured superpixel sets from image-level labels [29]. A weakly supervised image segmentation model, focusing on learning the semantic associations between superpixel sets, is proposed in [33].

However, for images with stripe texture as shown in Fig. 1 or lattice fence pattern which are captured for security as shown in Fig. 7, existing methods cannot achieve satisfied results because the stripe texture are more salient than the object edges, which makes these methods difficult to differentiate main structure from texture details.

To address the striped-texture image segmentation, we propose a novel framework consisting of three steps: (1)using frequency based band rejection filter to weaken stripes, since stripes in spatial domain correspond to an elongated pattern in frequency domain [13];
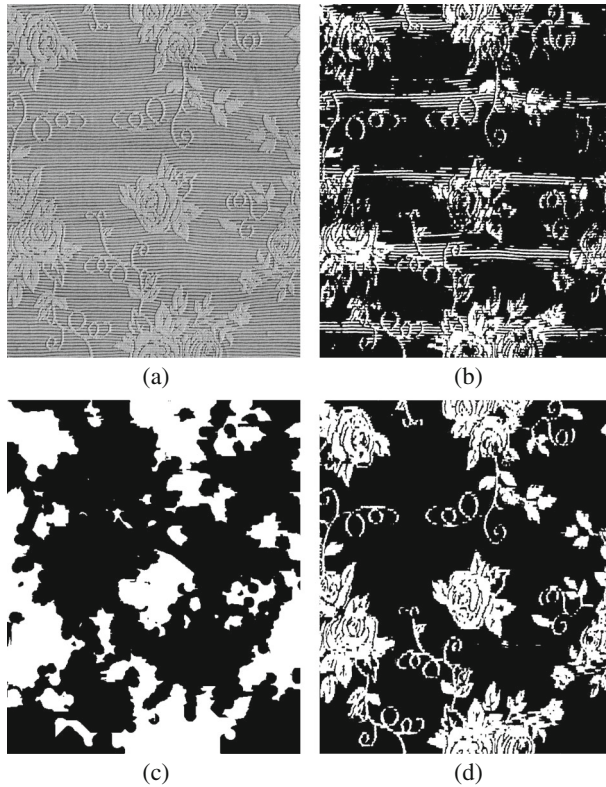
Fig. 1 **a** Input image; (**b**) result obtained by image smoothing and threshold segmentation; (**c**) result obtained by quick selection tool in Photoshop manually for half an hour; (**d**) result produced by our novel segmentation approach

(2)taking advantage of structure-preserving image smoothing [22] to remove the noise and highlight the main structures; (3)making use of an effective threshold method to classify pixels as foreground or background. As demonstrated in Fig. 1, the proposed method achieves a high-precision segmentation result. High quality image segmentation greatly improves the results of many applications. We present several fields that our method could benefit in the experiments part.

Our main contributions

– a framework of image segmentation for images with periodic stripe or lattice fence pattern, which benefit multimedia applications
– taking use of structure-preserving image smoothing technique for image segmentation.

2 Related work

As far as we know, it is the first work to combine image de-striping by frequency based filtering and structure-preserving image smoothing for striped-texture image segmentation.

De-striping and structure-preserving image smoothing are basic tools for remote sensing imagery(SRI) processing and image enhancement respectively. They have been addressed by a variety of works.

2.1 Image de-striping

Stripe noise in remote sensing images not only sharply degrades the image quality in the visual effect, but also risks their suitability for subsequent processing [2]. To remove the stripe and improve image quality, the de-striping methods have been proposed. From the viewpoints of methodology, existing stripe removal methods can be classified into three categories. The filtering-based methods are widely used. Pande-Chhetri et al. develop an image de-striping method based on wavelet analysis with fast and impressive de-striping results [13]. Münch et al. use the combination of wavelet and Fourier analysis to remove horizontal or vertical stripes [12]. Chen et al. propose to use a power finite-impulse response filter to remove the striping-induced frequency components [4].

For the second category method, the main idea is to rectify the distribution of the stripes to a reference distribution. Wegener proposes to calculate the histograms of stripe lines, and then match them to the reference [21]. Rakwatin et al. combine histogram matching with facet filter to reduce stripe noise in Moderate Resolution Imaging Spectroradiometer (MODIS) data [14].

The variational destriping methods regard the stripe removal problem as an ill-posed inverse problem. Shen et al. propose the Huber-Markov variational model to remove the stripes with spatial local adaptive edge-preserving ability [16]. Chang et al. treat the image and stripe component equally, and convert the image de-striping task as an image decomposition problem [23].

These works are either time-consuming or non-automatic. In this paper, we propose a filtering-based methods method that is automatic, effective and able to handle stripes in any direction.

2.2 Structure-preserving image smoothing

Structure-preserving image smoothing aims to extract main structures of images while removing texture details. There are two types of method.

One contains the optimization based filters. Xu et al. develop a robust method on new local variation measures to separate structure from texture [22]. Karacan et al. use region covariance matrices to capture local structure and texture information [9].

Another kind of edge-preserving smoothing techniques is weighted average based smoothing, which smooths an input image via weighted average affinities between neighboring pixel pairs. Zhang et al. introduce a scale-aware filter called as rolling guidance filter (RGF) [34]. RGF iteratively performs Gaussian filtering to remove small structures and joint bilateral filtering to recover edges. Zhang et al. design a new edge-aware structure, named segment graph, to represent the image and further develop a novel double weighted average image filter (SGF) based on the segment graph [24].

However, these methods mentioned above cannot handle images with stripes or lattice fence pattern because they take stripes and lattice fences as structures instead of texture.

Our approach combining de-striping with structure-preserving image smoothing achieves promising results on the test image dataset.
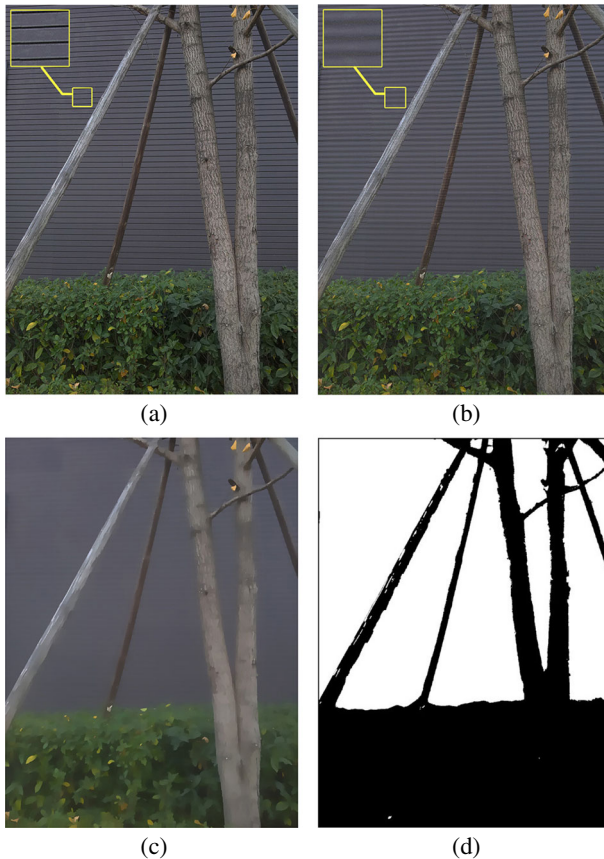
3 Approach

Our method has three distinct but interrelated stages- 1)Weakening stripes by frequency based filtering 2)Removing texture detail and extracting main structures 3)Producing segmentation results. Figure 2 demonstrates the procedure.

3.1 Frequency based filtering

Fourier transform is an important tool for image processing. For an image I(x,y) of size $M \times N$, its 2D discrete Fourier transform is defined as

$$F(u, v) = \sum_{x=0}^{M} \sum_{y=0}^{N} I(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} \tag{1}$$

(a)　　　　　　　　　(b)

(c)　　　　　　　　　(d)

Fig. 2 The procedure of our novel segmentation approach is shown step by step through an example. **a** Input image. **b** Image with weakened stripes. **c** The smoothed image. **d** Our finally segmentation result. In each top left part of (**a**) and (**b**), image medium scale details have been boosted, and the stripes are significantly weakened

According to [13], the striping patterns in the original image will be captured in the frequency domain as an elongated pattern in the direction perpendicular to the stripes. For example, horizontal stripes (Fig. 1a) are presented as a vertical central narrow band in Fourier domain (and contrarily, vertical ones as horizontal narrow band). Therefore, to weaken stripe texture, we apply a band rejection filter on the image spectrum.

Then the filtering process in frequency domain is showed as equation:

$$M(u, v) = F(u, v)H(u, v) \qquad (2)$$

in which $H(u, v)$ denotes the frequency filter and $M(u, v)$ is the filtered fourier power spectrum of the original image.

The de-striping method consists of three steps as shown in Fig. 3.

1. Get the power spectrum of the original image by 2D FFT.
2. Use a band rejection filter on the spectrum to mask frequency components which cause stripes.
3. Apply inverse 2D FFT to the masked spectrum.

It involves two key components for designing the band rejection filter.

(1) Locate the stripe frequency components We propose an effective method to detect which frequencies should be suppressed by the band rejection filter. Since stripes in spatial domain are reflected by narrow bands of high amplitude values in a direction orthogonal to the stripes [13], we accumulate the intensity values of the image spectrum along the narrow band direction to get the discrete accumulation curve S. The narrow bands of the most likely stripe frequencies are detected by finding the highest values of T, especially by looking for local maxima. We define the bands as stripe-band. The peaks show where the rejection filter should be applied.

For the image with horizontal stripes in Fig. 1a, its accumulation curve in spectrum (Fig. 4b) is shown in Fig. 4a. The red points imply three bands of frequencies along vertical axis to be rejected. Our method is not limited to vertical or horizontal stripes.

(2) Design the band rejection filter Based on the fact that the low frequencies in the Fourier transform correspond to the smooth areas of an image, we should separate the low frequencies from the stripe frequencies band detected above so as to preserve the image information while weakening the stripes. According to [35], we extend stripe-band by width W and project the values of the new band to the vertical line to get another accumulation curve T by (3).

$$T(v) = \sum_{w=1}^{W} F(u, v) \qquad (3)$$
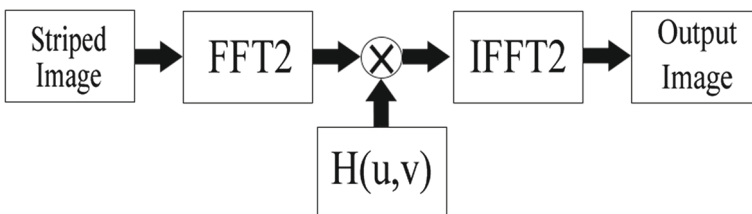
Fig. 3 Three steps of de-striping method. $H(u, v)$ denotes the frequency filter
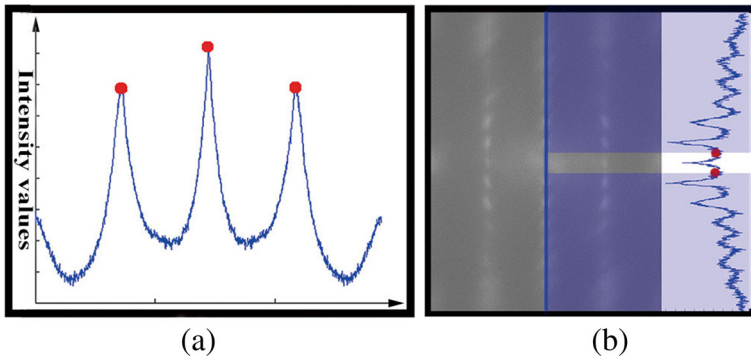
Fig. 4 a Discrete accumulation curve *S*. *Horizontal axis* is in direction of stripes. The three *red points* represent three places we apply band rejection filter. **b** Band rejection filter. The *bluish line* is the extended band to get another accumulation curve and the region between two red points is separated from the rejection band

in which $F(u, v)$ is the fourier power spectrum of the original image, and W represents the width of narrow bands of high amplitude values in the spectrum.

By finding local extreme values, we look for two points that separate the low frequencies in the center of the detected frequencies band. Then the region between the two points will be separated from the rejection band area. We treat each band the same way. In Fig. 4, the region between two points along central band is separated from the band.

We apply the rejection filter to the spectrum to weaken the stripes, and then restore the image. An example is shown in Fig. 2a. As we see in the boosted image details, stripes are significantly weakened.

In the case of lattice fence image, it can be treated as two stripes in different directions. We apply the method mentioned above to image in both directions to weaken lattice fence texture.

3.2 Image smoothing

By filtering the images in the frequency domain, we remove the stripe or fence texture. In this de-striping process, noise is introduced, which makes segmentation more challenging. Using the images obtained above as input, we apply structure-preserving image smoothing method on filtered image to smooth image and extract the main structure of the image. In our paper, we use relative total variation(RTV) model to perform image smoothing [22].

3.2.1 Relative total variation (RTV) model

Relative total variation(RTV) model is proposed to capture the nature of structure and texture which achieves promising results [22].

RTV model contains a general pixel-wise windowed total variation measure, written as

$$
\begin{aligned}
D_x(p) &= \sum_{q \in R(p)} g_{p,q} \cdot \left| (\partial_x S)_q \right| \\
D_y(p) &= \sum_{q \in R(p)} g_{p,q} \cdot \left| (\partial_y S)_q \right|
\end{aligned}
\tag{4}
$$

where q belongs to $R(p)$ and $R(p)$ is the rectangular region centered at pixel p. $D_x(p)$ and $D_y(p)$ are windowed total variations in the x and y directions for pixel p, which count the

absolute spatial difference within the window $R(p)$. $g_{p,q}$ is a weighting function defined according to spatial affinity, expressed as

$$g_{p,q} \propto exp(-\frac{(x_p - x_q)^2 - (y_p - y_q)^2}{2\sigma^2}) \tag{5}$$

where σ controls the spatial scale of the window.

To help distinguish prominent structures from the texture elements, besides D, RTV model also contains a windowed inherent variation, expressed as

$$\begin{aligned} L_x(p) &= |\sum_{q \in R(p)} g_{p,q} \cdot (\partial_x S)_q| \\ L_y(p) &= |\sum_{q \in R(p)} g_{p,q} \cdot (\partial_y S)_q| \end{aligned} \tag{6}$$

L captures the overall spatial variation.

The objective function is finally expressed as

$$\underset{S}{argmin} \sum_p (S_p - I_p)^2 + \lambda \cdot (\frac{D_x(p)}{L_x(p) + \varepsilon} + \frac{D_y(p)}{L_y(p) + \varepsilon}) \tag{7}$$

Thanks to the proposed relative total variation measure, the RTV model makes main structures of images easily distinguished from detail texture and achieves promising results of structure-preserving image smoothing.

3.2.2 Examples

Figure 5b shows the result of the relative total variation model used directly on a striped texture image. The reason why the performance is poor is that structure-preserving images smoothing implicitly assumes that salient edges only come from object contours/boundaries which is not true for stripe images. In this model, parameter σ represents the size of the texture elements. When σ is set large, the stripes are considered as large texture elements, but this will result in more blurred images.

Frequency-domain filtering makes stripes less obvious, the RTV model considers these not as main structures but as texture details. Comparison results are shown in Fig. 5b and d. Parameters, e.g. σ, λ(smoothness), of two approaches for each image are set to be the same values. In Fig. 5b, the stripes are not smoothed, while they are smoothed in Fig. 5d.

Figure 6 shows the comparison of results of our method and method without lattice fence weakening on a lattice fence pattern image.
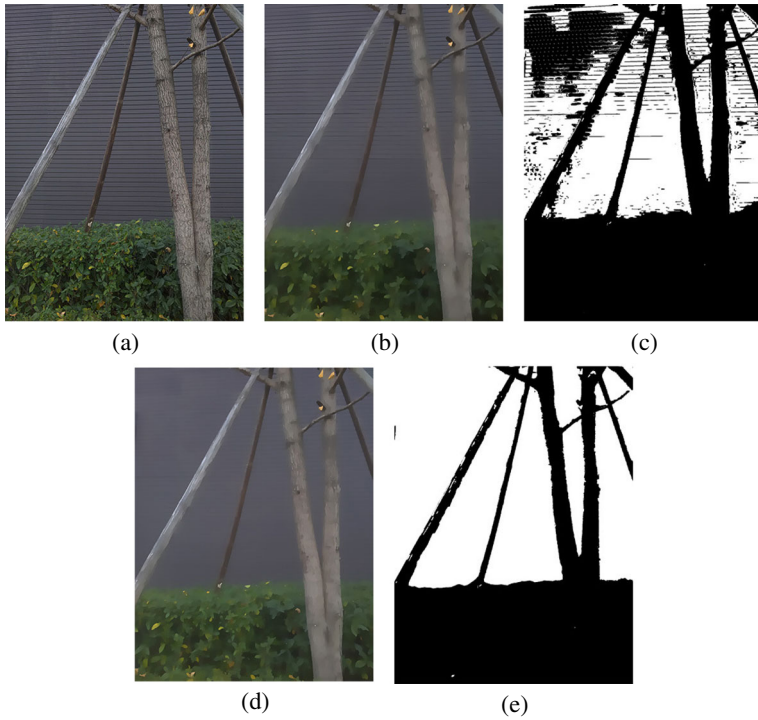
3.3 Segmentation

After structure-preserving image smoothing, smoothed images contain a few intensity values and the main structures are extracted. In the third step, we employ an effective threshold method for segmentation. In Figs. 5c and e, 6c and e, segmentation results of two approaches are presented. Our approach achieves superior results.

4 Experimental results

Our Method is applied on a variety of test images, including fence image of prison and textile dataset which contains images with stripe texture obtained from a textile mill. Very promising results are achieved.

(a)   (b)   (c)

(d)   (e)

Fig. 5 Comparison of a striped texture image. **a** Input images. **b** Relative total variation model. **c** Segmentation results after RTV model without stripes weakening. **d** Smoothed images by our approach. **e** Segmentation results obtained by our approach

(a)   (b)   (c)

(d)   (e)
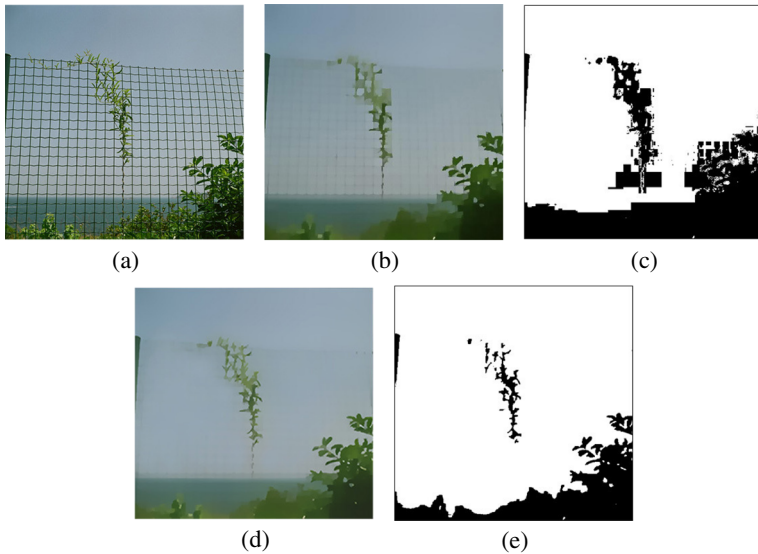
Fig. 6 Comparison of a lattice fence image. **a** Input images. **b** Relative total variation model. **c** Segmentation results after RTV model without lattice fence weakening. **d** Smoothed images by our approach. **e** Segmentation results obtained by our approach

Fig. 7 **a** Input fence image. **b** Segmentation results

As shown in Fig. 7, we have achieved great performance of image segmentation on prison image occluded by lattice fence. It is known to all that image segmentation plays an important role in surveillance security area such as train stations and prisons. But it remains challenging to detect and segment meaningful objects such as human out of cluttered background, especially suffering from stripe texture noise and occlusion by lattice fence. Result in Fig. 7 shows that our approach works well in these situations, which has great value in public security area such as prison surveillance and crime scene investigation.

Figure 8a shows a textile image with background formed by stripe texture. The structure-preserving image smoothing algorithms can not smooth the stripes directly because large gradients of stripes make them mistaken as structures. As we can see, our method separates the flower pattern completely from the stripes in Fig. 8b. Our approach is valuable in textile industry. Based on the satisfied segmentation result of stripe jacquard fabric image, we could combine textile printing with jacquard weave to produce new textile products which are both colorful and high-grade. First, we get image of jacquard fabric and Segment the flower pattern. Second, we register the designed color pattern with images of deformed jacquard fabric. It will promote textile industry.
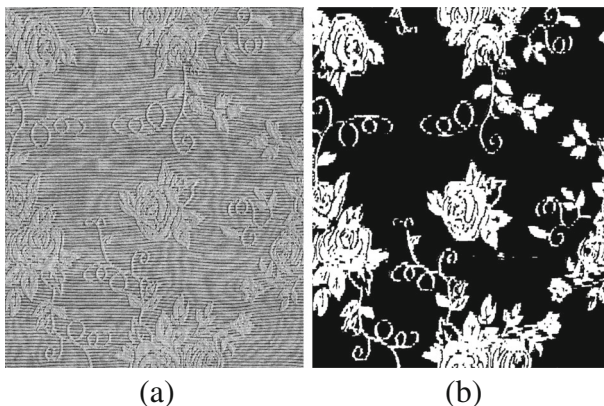
Fig. 8 **a** Input fabric image. **b** Segmentation results

Fig. 9 **a** An image with characters. **b** Segmentation results

Our approach could benefit other application. Figure 9a shows an image of characters embedded on fence. With our method applied to the image, the characters are successfully extracted as shown in Fig. 9b. The optical character recognition(OCR) in natural scenes of characters embedded on fence benefits from our new approach.

5 Conclusions and future work

We have presented an image segmentation method for images with stripe or fence texture, with applications to multimedia security. Our approach consists of three steps: de-striping by frequency filtering, structure-preserving image smoothing and classifying pixels as foreground and background. Very promising results have been achieved by our method. In future work, we will focus on image segmentation with repeated patterns not restricted to stripe patterns, which could benefit more multimedia applications.

References

1. Bao B-K, Liu G, Hong R, Yan S, Changsheng X (2013) General subspace learning with corrupted training data via graph embedding. IEEE Trans Image Process 22(11):4380–4393
2. Bouali M, Ladjal S (2011) Toward optimal destriping of modis data using a unidirectional variational model. IEEE Trans Geosci Remote Sens 49(8):2924–2935
3. Boykov Y, Jolly M-P (2001) Interactive graph cuts for optimal boundary and region segmentation of objects in nd images. In: Proceedings of Eighth IEEE international conference on computer vision. Vancouver, pp 105–112
4. Chen J, Shao Y, Guo H, Wang W, Zhu B (2003) Destriping cmodis data by power filtering. IEEE Trans Geosci Remote Sens 41(9):2119–2124
5. Comaniciu D, Meer P (2002) Mean shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal Mach Intell 24(5):603–619
6. Fang M-Y, Kuan Y-H, Kuo C-M (2012) Effective image retrieval techniques based on novel salient region segmentation and relevance feedback. Multimed Tools Appl 57:501–525
7. Grady L, Singh V, Kohlberger T, Alcino C, Bahlmann C (2012) Automatic segmentation of unknown objects, with application to baggage security. In: Proceedings of IEEE European conference on computer vision. Firenze, pp 430–444

8. Hong C, Zhu J, Jun Y, Cheng J, Chen X (2014) Realtime and robust object matching with a large number of templates. Multimed Tools Appl 75:1459–1480
9. Karacan L, Erdem E, Erdem A (2013) Structure-preserving image smoothing via region covariances. ACM Trans Graph 32:6
10. Li T, Chang H, Wang M, Ni B, Hong R, Yan S (2015) Crowded scene analysis: a survey. IEEE Trans Circ Syst Video Technol 25(3):367–386
11. Lin C, Pun C-M, Huang G (2016) Highly non-rigid video object tracking using segment-based object candidates. Multimed Tools Appl
12. Mnch B, Trtik P, Marone F, Stampanoni M (2009) Stripe and ring artifact removal with combined wavelet - fourier filtering. Opt Express 17(10):8567–8591
13. Pande-Chhetri R, Abd-Elrahman A (2011) De-striping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering. ISPRS J Photogramm Remote Sens 66(5):620–636
14. Rakwatin P, Takeuchi W, Yasuoka Y (2009) Restoration of aqua modis band 6 using histogram matching and local least squares fitting. IEEE Trans Geosci Remote Sens 47(2):613–627
15. Sang J, Changsheng X, Liu J (2012) User-aware image tag refinement via ternary semantic analysis. IEEE Trans Multimed 14(3-2):883–895
16. Shen H, Zhang L (2009) A map-based algorithm for destriping and inpainting of remotely sensed images. IEEE Trans Geosci Remote Sens 47(5):1492–1502
17. Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22(8):888–905
18. Tang J, Tao D, Qi G-J, Huet B (2014) Social media mining and knowledge discovery. Multimed Syst 20(6):633–634
19. Vese L, Chan T (2002) A multiphase level set framework for image segmentation using the mumford and shah model. Int J Comput Vis 50(3):271–293
20. Wang W, Yan Y, Zhang L, Hong R, Sebe N (2016) Collaborative sparse coding for multi-view action recognition. IEEE Multimed Mag 23(4):80–87
21. Wegener M (1990) Destriping multiple sensor imagery by improved histogram matching. Int J Remote Sens 11(5):859–875
22. Xu L, Yan Q, Xia Y, Jia J (2012) Structure extraction from texture via relative total variation. ACM Trans Graph 31:6
23. Yi C, Yan L, Tao W, Zhong S (2016) Remote sensing image stripe noise removal: from image decomposition perspective. IEEE Trans Geosci Remote Sens 54(12):7018–7031
24. Zhang F, Dai L, Xiang S, Zhang X (2015) Segment graph based image filtering: fast structure-preserving smoothing. In: Proceedings of IEEE conference on computer vision and pattern recognition. Boston, pp 361–369
25. Zhang L, Hong R, Gao Y, Ji R, Dai Q, Li X (2016) Image categorization by learning a propagated graphlet path. IEEE Trans Neural Netw Learn Syst 27(3):674–685
26. Zhang L, Li X, Nie L, Yan Y, Zimmermann R (2016) Semantic photo retargeting under noisy image labels. ACM Trans Multimed Comput Commun Appl 12(3):37
27. Zhang L, Li X, Nie L, Yi Y, Xia Y (2016) Weakly supervised human fixations prediction. IEEE Trans CyBern 46(1):258–269
28. Zhang L, Song M, Li N, Bu J, Chen C (2009) Feature selection for fast speech emotion recognition. In: Proceedings of ACM international conference on multimedia. Beijing, pp 753–756
29. Zhang L, Song M, Liu Z, Liu X, Bu J, Chen C (2013) Probabilistic graphlet cut: exploring spatial structure cue for weakly supervised image segmentation. In: Proceedings of IEEE conference on computer vision and pattern recognition. Portland, pp 1908–1915
30. Zhang L, Song M, Qi Z, Liu X, Jiajun B, Chen C (2013) Probabilistic graphlet transfer for photo cropping. IEEE Trans Image Process 21(5):2887–2897
31. Zhang L, Wang M, Hong R, Yin B-C, Li X (2016) Large-scale aerial image categorization using a multitask topological codebook. IEEE Trans CyBern 46(2):535–545
32. Zhang L, Yang Y, Wang M, Hong R, Nie L, Li X (2016) Detecting densely distributed graph patterns for fine-grained image categorization. IEEE Trans Image Process 25(2):553–565
33. Zhang L, Yi Y, Gao Y, Wang C, Yi Y, Li X (2014) A probabilistic associative model for segmenting weakly-supervised images. IEEE Trans Image Process 23(9):4150–4159
34. Zhang Q, Shen X, Xu L, Jia J (2014) Rolling guidance filter. In: Proceedings of European conference on computer vision. Zurich, pp 815–830
35. Zhang Z, Shi Z, Guo W, Huang S (2005) Adaptively image de-striping through frequency filtering. In: ICO20: Opt. Inf. Proc., Proc. SPIE: 6027, pp 989–996

Jinxiong Ren is currently working toward the Ph.D. degree at the College of Computer Science, Zhejiang University, Hangzhou, China. His research area includes image segmentation, image registration, point matching, and deep learning.
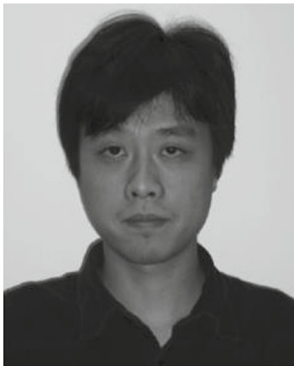
Gang Chen received the Ph.D. degree in computer science from Zhejiang University, Hangzhou, China. He is a Professor with the College of Computer Science and the Director of the Database Lab, Zhejiang University, Hangzhou, China. He has successfully led investigations in research projects that aim at building China's indigenous database management systems. His research interests range from relational database systems to large-scale data management technologies supporting massive Internet users. Dr. Chen is a member of the ACM and a senior member of China Computer Federation.

Xiaoyan Li received the B.S. and Ph.D. degrees in computer science from Zhejiang University in 2004 and 2009 respectively. From Jun. 2006 to Dec. 2006, she was a visiting student at Database Lab in SOC, National University of Singapore. Her research area includes multimedia database, semantic learning and modeling, Web image retrieval.

Kuang Mao received the Ph.D. degree in computer science from Zhejiang University, Hangzhou, China. From 2013 to 2014, he was a Research Intern with the SESAME Group at the National University of Singapore, Singapore. His research area includes recommendation systems, singing song recommendation, graph ranking algorithms, and probabilistic modeling.