A multiphase texture segmentation method based on local intensity distribution and Potts model*

WANG Jing (王靖)[1,2]**, **ZHENG Yong-guo** (郑永果)[1], **PAN Zhen-kuan** (潘振宽)[2], **ZHANG Wei-zhong** (张维忠)[2], **and WANG Guo-dong** (王国栋)[2]

1. College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China

2. College of Information Engineering, Qingdao University, Qingdao 266071, China

(Received 2 March 2015; Revised 14 April 2015)

Because texture images cannot be directly processed by the gray level information of individual pixel, we propose a new texture descriptor which reflects the intensity distribution of the patch centered at each pixel. Then the general multiphase image segmentation model of Potts model is extended for texture segmentation by adding the region information of the texture descriptor. A fast numerical scheme based on the split Bregman method is designed to speed up the computational process. The algorithm is efficient, and both the texture descriptor and the characteristic functions can be implemented easily. Experiments using synthetic texture images, real natural scene images and synthetic aperture radar images are presented to give qualitative comparisons between our method and other state-of-the-art techniques. The results show that our method can accurately segment object regions and is competitive compared with other methods especially in segmenting natural images.

Document code: A **Article ID:** 1673-1905(2015)04-0307-6

DOI 10.1007/s11801-015-5036-8

Different approaches deal with the extraction of homogeneous features from textures. A very popular class of texture feature extraction is the transformation-based feature as the structure tensor[1] or the Gabor filter[2]. Another class of method for texture feature extraction is the statistics-based one as the co-occurrence matrix[3]. Besides the two classes of feature extraction methods, other texture descriptors have been proposed. Semi-local region texture descriptor[4] combined Beltrami framework representation and semi-local image information. It is equivalent to smoothing the gradient image with the square kernel[5]. The local binary pattern (LBP) operator, introduced by Ojala et al[6], encodes the local microstructure between the image pixels, and a histogram of LBP can be built to achieve a description of the texture feature[7]. One of the latest improved versions of LBP is concave-convex LBP[8], which is calculated after the neighborhoods of the image are divided into concave and convex categories.

Many models and algorithms have been proposed to perform the texture segmentation. Ref.[4] used the popular Kullback-Leibler distance to design an active contour model and introduced a segmentation algorithm based on the split Bregman method. Ref.[5] improved edge-weighted centroidal voronoi tessellation and proposed a narrow banding algorithm to accelerate the implementation. Ref.[9] coupled the geodesic active contour with the multiphase successive active contour model and approximately solved it by a multilayer graph method. Ref.[10] proposed a multiphase fuzzy region competition model for texture image segmentation, in which each region is represented by a fuzzy membership function and a probability density function. In Ref.[11], an improved active contour model was proposed by using the non-local information in multi-scale texture feature space.

In this paper, we propose a method for multiphase texture segmentation by minimizing the segmentation energy with some constraint conditions. The new data term in the proposed model substitutes the mean value and the Euclidean distance for the complex mathematical operations, such as the logarithmic function and Gaussian function used in Ref.[10]. Compared with two phase segmentation models[11], our multiphase segmentation model has broader application prospects. While compared with other multiphase model, for example that in Ref.[9], our model has more concise representation to describe each phase. Furthermore, our method does not rely on a training stage and is unsupervised. A new and simple texture descriptor is used in our model to segment complex textural images, which helps to speed up the segmentation. These properties make our method applicable to a wide range of texture segmentation problems.

* This work has been supported by the National Natural Science Foundation of China (No.61170106).

** E-mail: cfyljnan@sohu.com

Although the intensity distribution of single pixel is disorganized in texture images, we believe that it has some regularity when thinking of a small pixel patch. Accordingly, the local intensity distribution probability is defined to represent texture units localized in image neighborhoods. Considering a local neighborhood with $P\times P$ pixels, the central pixel is (x_c, y_c). Let $\boldsymbol{R}$ be the joint distribution of the gray levels in this neighborhood, which can be expressed as

$$\boldsymbol{R}=\boldsymbol{r}(g_1,g_2,...,g_{P\times P}), \tag{1}$$

where g_i (i=1, …, $P\times P$) represents the gray level of pixel in the local neighborhood centered at (x_c, y_c).

To facilitate the statistics of gray level distribution, we sort the 256 gray levels into 26 intensity bins. That is, gray levels from 0 to 9 belong to the first bin b_1, gray levels from 10 to 19 belong to the second bin b_2, and so forth. Then much of the information in the original joint gray level distribution expressed as Eq.(1) about the textural characteristics is conveyed by the joint intensity bin distribution:

$$\boldsymbol{R}\approx\boldsymbol{r}(ib_1,ib_2,...,ib_{P\times P}), \tag{2}$$

where ib_i (i=1, …, $P\times P$) represents the intensity bin that pixel i belongs to.

The local intensity distribution probability for such a $P\times P$ neighborhood is estimated as follows. Initialize integer variables n_i (i=1, …, 26) to be 0. And check every pixel in the patch. If its gray level belongs to the intensity bin b_i, 1 is added to n_i. After all the pixels are calculated, n_i records the number of pixels whose gray level is between $(i-1)\times10$ and $(i\times10-1)$. Let $d_i=\dfrac{n_i}{P\times P}(i=1,...,26)$. The 26-dimensional vector **lidp**$(x_c, y_c)=(d_1, d_2, d_3,…, d_{26})$ is exactly the local intensity distribution probability of pixel (x_c, y_c). It records the occurrences of various intensity values in the neighborhood of each pixel in a vector, and is a highly discriminative texture operator.

The Potts model is derived from statistical mechanics, and provides a uniform framework for multiphase image segmentation. If $f(x, y)$: $\Omega\to R$ defines the image intensity of the segmented region Ω, the Potts model for multiphase image segmentation can be formulated as the following energy minimization problem:

$$\underset{\{c_i\}_{i=1}^n}{\mathrm{Min}}\ \underset{\{\Omega_i\}_{i=1}^n}{\mathrm{Min}}\left\{\sum_{i=1}^n\alpha_i\int_{\Omega_i}Q_i(c_i,x,y)\mathrm{d}x\mathrm{d}y+\sum_{i=1}^n\gamma_i|\partial\Omega_i|\right\},$$

$$\text{s. t. }\Omega=\bigcup_{i=1}^n\Omega_i,\quad\Omega_i\bigcap_{i\neq j}\Omega_j=\varnothing, \tag{3}$$

where $|\partial\Omega_i|$ expresses the boundary length or surface area of subregion Ω_i, $Q_i(c_i, x, y)$ is the estimation function of characteristic parameter c_i in subregion Ω_i, and α_i and γ_i are penalty parameters of the estimation function term and length term, respectively.

Under the variational framework, subregions are identified with characteristic functions. According to different design methods for characteristic functions, the model expressed as Eq.(3) can be transformed to variational level set model, discrete labeling function model and Gamma-convergence model. In this paper, we make use of the second design method. That is, several binary labeling functions act as the characteristic functions. We define $\phi_i(x,y)=\begin{cases}1, & \text{if } (x,y)\in\Omega_i\\ 0, & \text{otherwise}\end{cases}$ for each phase, and the constraint condition of Eq.(3) requires $\sum_{i=1}^n\phi_i(x,y)=1$. Then, Eq.(3) can be stated as the following computable energy minimization problem:

$$\underset{\{c_i\}_{i=1}^n}{\mathrm{Min}}\ \underset{\{\phi_i\}_{i=1}^n}{\mathrm{Min}}\left\{\sum_{i=1}^n\alpha_i\int_\Omega Q_i(c_i,x,y)\phi_i\mathrm{d}x\mathrm{d}y+\sum_{i=1}^n\gamma_i\int_\Omega|\nabla\phi_i|\mathrm{d}x\mathrm{d}y\right\},$$

$$\text{s.t. }\phi_i\in\{0,1\},\ \sum_{i=1}^n\phi_i=1\ . \tag{4}$$

When a piecewise constant image[12] is segmented, the form of estimation function is $Q_i(c_i, x, y)=[c_i-f(x, y)]^2$, where c_i is the average intensity in subregion Ω_i. As to the piecewise smooth image[13], the estimation function is $Q_i(c_i,x,y)=[c_i-f(x,y)]^2+\rho_i|\nabla c_i|^2$, where ρ_i is the penalty parameter.

When the textures have the same intensity, it is very difficult for the standard segmentation model expressed as Eq.(4) to tell them apart. So we need to change the standard segmentation model to allow the computer to segment the textured image.

The notion of texture is undefined at single pixel level, and it is always associated with some set of pixels[14]. This motivates us to replace the single point information quantified by $f(x, y)$ with the regional information captured by the local intensity distribution probability **lidp**(x, y) estimated by all pixels centered at the pixel (x, y) which belong to a $P\times P$ neighborhood.

We use squared differences of the local intensity distribution probability **lidp**(x, y) and the average local intensity distribution probability in each subregion as a similarity measure. The formulation of the Potts model that uses local regional information is proposed as

$$\underset{\{c_i\}_{i=1}^n}{\mathrm{Min}}\ \underset{\{\phi_i\}_{i=1}^n}{\mathrm{Min}}\left\{\sum_{i=1}^n\alpha_i\int_\Omega[\frac{1}{b}\sum_{j=1}^b\lambda_j(\mathrm{lidp}(x,y,j)-c_i^j)^2]\phi_i\mathrm{d}x\mathrm{d}y+\sum_{i=1}^n\gamma_i\int_\Omega|\nabla\phi_i|\mathrm{d}x\mathrm{d}y\right\},$$

$$\text{s.t. }\phi_i\in\{0,1\},\ \sum_{i=1}^n\phi_i=1\ , \tag{5}$$

where b is the number of intensity bins, equal to 26 here. The positive scalars λ_j (j=1,…, 26) are weight parameters for the jth bin. lidp(x, y, j) is the jth component of the local intensity distribution probability **lidp**(x, y). And c_i^j is the average local intensity distribution probability of the jth bin in subregion Ω_i.

For the constraint condition of Eq.(5), we can use penalty function method to reformulate the model expressed as Eq.(5) as follows:

$$\min_{\{c_i\}_{i=1}^n}\min_{\{\phi_i\}_{i=1}^n}\left\{\sum_{i=1}^{n}\alpha_i\int_\Omega\left[\frac{1}{b}\sum_{j=1}^{b}\lambda_j(\mathrm{lidp}(x,y,j)-c_i^j)^2\right]\phi_i\mathrm{d}x\mathrm{d}y+\right.$$

$$\left.\sum_{i=1}^{n}\gamma_i\int_\Omega|\nabla\phi_i|\mathrm{d}x\mathrm{d}y+\frac{\mu}{2}\int_\Omega(\sum_{i=1}^{n}\phi_i-1)^2\mathrm{d}x\mathrm{d}y\right\},$$

$$\text{s.t.}\quad \phi_i\in\{0,1\}, \tag{6}$$

where μ is a penalty parameter.

Eq.(6) is a multivariate minimization problem, it is usually solved via alternative optimization procedure, i.e., fixing ϕ_i to solve the minimization with respect to c_i^j, and then vice versa. When ϕ_i is fixed, it can be obtained as

$$c_i^j=\frac{\int_\Omega\mathrm{lidp}(x,y,j)\phi_i(x,y)\mathrm{d}x\mathrm{d}y}{\int_\Omega\phi_i(x,y)\mathrm{d}x\mathrm{d}y},(i=1,2,...,n). \tag{7}$$

Next, we design split Bregman method[15] to calculate the binary characteristic functions ϕ_i (i=1, 2,…, n). For the constraint condition of Eq.(6), we relax $\phi_i\in\{0,1\}$ to $\phi_i\in[0,1]$. n auxiliary variables $\boldsymbol{w}_i$ (i=1, 2,…, n) are introduced, so that $\boldsymbol{w}_i\approx\nabla\phi_i$ when the energy function achieves the minimum. And n Bregman iterative parameters $\boldsymbol{b}_i$ (i=1, 2,…, n) are introduced to reduce the dependence of the penalty parameter. Thus, Eq.(6) can be transformed into the following iterative minimization formulation as

$$(\phi_i^{k+1},\boldsymbol{w}_i^{k+1})=$$

$$\mathrm{Arg}\min_{\boldsymbol{w}_i,\phi\in[0,1]}\left\{\sum_{i=1}^{n}\alpha_i\int_\Omega\left[\frac{1}{b}\sum_{j=1}^{b}\lambda_j(\mathrm{lidp}(x,y,j)-c_i^j)^2\right]\phi_i\mathrm{d}x\mathrm{d}y+\right.$$

$$\sum_{i=1}^{n}\gamma_i\int_\Omega|\boldsymbol{w}_i|\mathrm{d}x\mathrm{d}y+\frac{\theta}{2}\sum_{i=1}^{n}\int_\Omega\left(\boldsymbol{w}_i-\nabla\phi_i-\boldsymbol{b}_i^{k+1}\right)^2\mathrm{d}x\mathrm{d}y+$$

$$\left.\frac{\mu}{2}\int_\Omega\left(\sum_{i=1}^{n}\phi_i-1\right)^2\mathrm{d}x\mathrm{d}y\right\}, \tag{8}$$

where θ>0 is a penalty parameter, $\boldsymbol{w}_i$ and $\boldsymbol{b}_i$ are vectors, and $\boldsymbol{b}_i$ is updated by $\boldsymbol{b}_i^{k+1}=\boldsymbol{b}_i^k+\nabla\phi_i^k-\boldsymbol{w}_i^k$ and $\boldsymbol{b}^0=\boldsymbol{w}^0=\boldsymbol{0}$. Using alternating minimization strategy, we can get the following Euler-Lagrange equation on ϕ_i as

$$\begin{cases}\alpha_i\left[\frac{1}{b}\sum_{j=1}^{b}\lambda_j(\mathrm{lidp}(x,y,j)-c_i^j)^2\right]+\\ \theta\nabla\cdot\left(\boldsymbol{w}_i^k-\nabla\phi_i-\boldsymbol{b}_i^{k+1}\right)+\mu\left(\sum_{j=1}^{n}\phi_j-1\right)=0,\ \text{in }\Omega.\\ \left(\boldsymbol{w}_i^k-\nabla\phi_i-\boldsymbol{b}_i^{k+1}\right)\cdot\boldsymbol{n}=0,\qquad\qquad\text{on }\partial\Omega\end{cases} \tag{9}$$

The results of Eq.(9) should be projected on [0, 1] with the formula

$$\tilde{\phi}_i=\mathrm{Max}(0,\mathrm{Min}(\phi_i,1)). \tag{10}$$

After projection, the binary characteristic function can be obtained by the following thresholding technique:

$$\overline{\phi}_i=\begin{cases}1, & \tilde{\phi}_i\ge\eta\\ 0, & \text{otherwise}\end{cases},\ \forall\eta\in(0,1). \tag{11}$$

The sub-optimization on $\boldsymbol{w}_i$ leads to the following generalized soft thresholding formula in analytical form as

$$\boldsymbol{w}_i^{k+1}=\mathrm{Max}\left(\left|\nabla\phi_i^{k+1}+\boldsymbol{b}_i^{k+1}\right|-\frac{\gamma_i}{\theta},0\right)\frac{\nabla\phi_i^{k+1}+\boldsymbol{b}_i^{k+1}}{\left|\nabla\phi_i^{k+1}+\boldsymbol{b}_i^{k+1}\right|}. \tag{12}$$

Experiments are performed to investigate the performance of the proposed model and algorithm on texture segmentation. The various tested images include synthetic texture images, natural images and synthetic aperture radar (SAR) images. The synthetic textural images shown in Fig.1(a) and (g) were generated from the Brodatz data set[16]. The natural textural images shown in Fig.5(a), (d), (g), (j) and (m) were taken from the Berkeley data set[17]. SAR images and the other images all appeared in the recent image segmentation literatures of Refs.[9,10,18-20].

There are several parameters which need to be tuned. The parameters α_i (i=1,…, n) are the coefficients of data terms. In most cases, we choose α_i to be around 0.01. γ_i (i=1,…, n) are the coefficients of smoothness terms and affect the boundary lengths of the segmented subregions. The general choice of γ_i is 1 000. Penalty parameters θ and μ are set to be 0.1 and 1.0, respectively. λ_j (j=1,…, b) control the weights of squared differences between b intensity bins and are always set to be 1.0 in all experiments. P is the side length of pixel patch used in the calculation of local intensity distribution probability. It should be bigger than the scale of the texture structure, so that enough discriminative information can be contained in the patch. For natural images used in this paper, P can be set to be 5, 7 or 9.

All experiments are performed using Matlab7.0.1 on a Windows 7 platform with an Intel Core 2 Duo CPU at 3.1 GHz and 4 GB memory. The stopping criterion is based on the relative error formula $|E^{k+1}-E^k|/E^k\le\varepsilon$, where ε is a small prescribed tolerance, and we set it as 10^{-2} in numerical experiments. We display the segmentation results by painting the contour on the original image. Our results are compared with those of other state-of-the-art segmentation methods.

Fig.1 shows three two-phase synthetic texture images and the segmentation results. Fig.1(d) consists of a chirp-like brick wall background and a Brodatz texture object. Note that the average gray values between the two textured regions in these texture images are not too different. Nevertheless, our method works quite well. Compared with Fig.1(c) obtained by the method in Ref.[4], our result shown in Fig.1(b) has a more accurate and smooth boundary. The other two segmentation results shown in

Fig.1(e) and (h) are competitive with Fig.1(f) and (i) obtained by other methods[10,21], respectively.

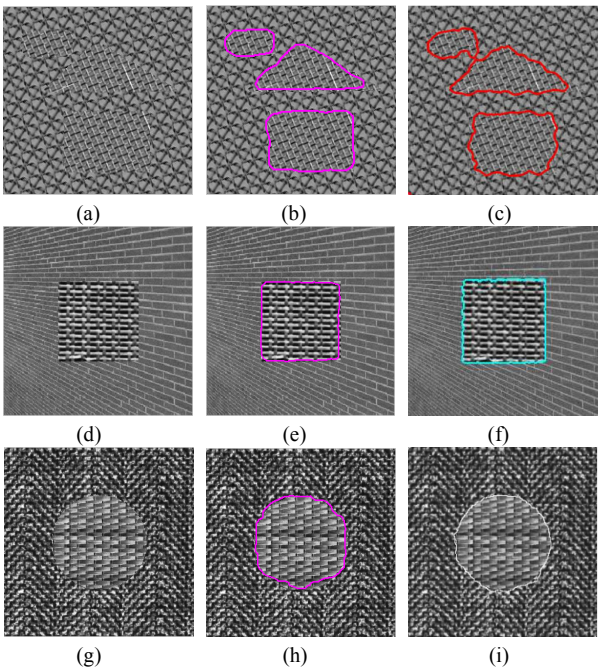


Fig.1 Two-phase segmentation results: (a)(d)(g) Original images; (b)(e)(h) Our results; (c) Result from Ref.[4]; (f) Result from Ref.[10]; (i) Result from Ref.[21]

In Figs.2–4, we test our method with 3-phase and 4-phase synthetic texture image segmentation. The segmentation results are still satisfactory. In the result of Fig.2(c) obtained by the method in Ref.[10], the left top corner of the square in the middle is a bit round. While in our result of Fig.2(b), the square boundary is of four sharp corners, which is more accurate. Fig.3(b) and Fig.4(b) are closer to their ground truth.

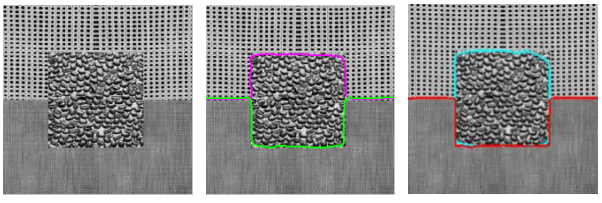


Fig.2 Three-phase segmentation results: (a) A synthetic image; (b) Our result; (c) Result from Ref.[10]

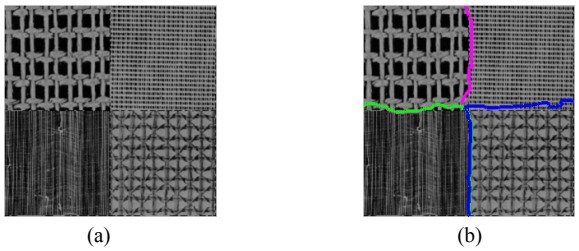


Fig.3 Four-phase segmentation result (I): (a) A synthetic image; (b) Our result

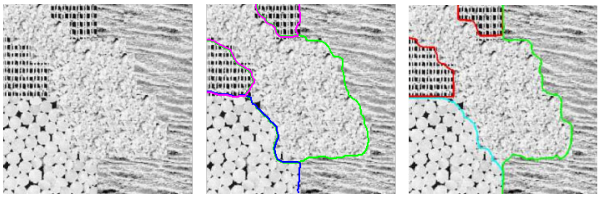
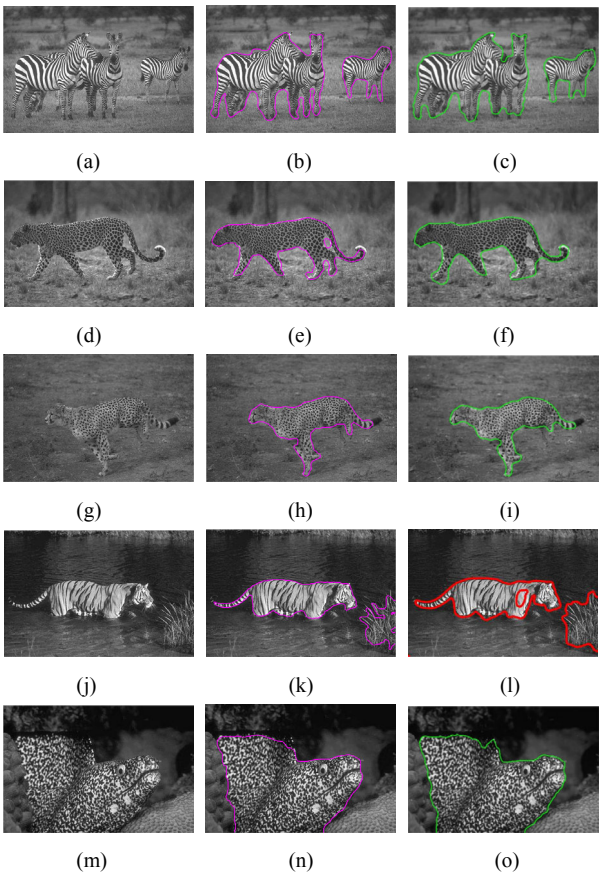
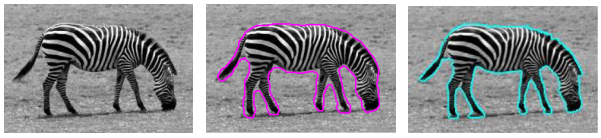


Fig.4 Four-phase segmentation result (II): (a) A synthetic image; (b) Our result; (c) Result from Ref.[10]

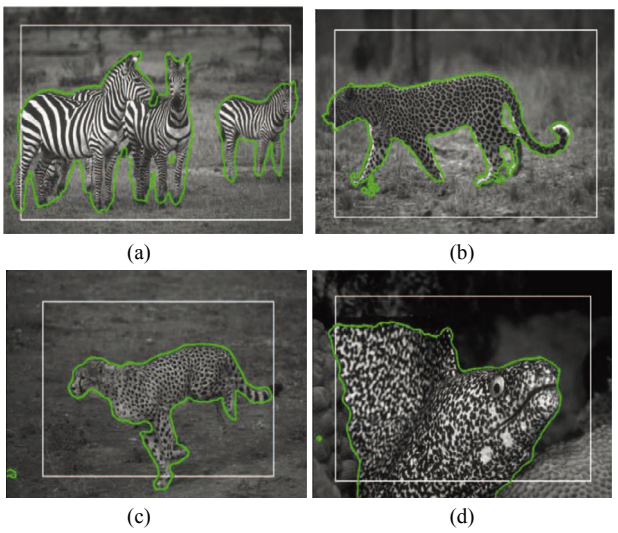
In Fig.5, six natural texture images are tested. Real-world images are more interesting, and can provide in-depth insight into the segmentation performance. For the images of the herd of zebras in Fig.5(a) and a walking leopard in Fig.5(d), our results in Fig.5(b) and Fig.5(e) are much better than Fig.5(c) and Fig.5(f) in the leg segmentation boundaries. In Fig.5(h), the segmented hind leg of a running leopard is longer than the corresponding part in Fig.5(i). Compared with Fig.5(l), the segmentation contour in Fig.5(k) is tighter to the tiger body, and there is no hole in the shoulder. Some latest results of these natural images are illustrated in Fig.6. They were taken from Ref.[11] which proposed a two-stage segmentation approach in multi-scale texture feature space. It's not difficult to see that our results in Fig.5 are superior to the results from Ref.[11] in Fig.6. The main advantages of ours are more adaption to the body curves of the animals and the absence of speckles.





(p) (q) (r)

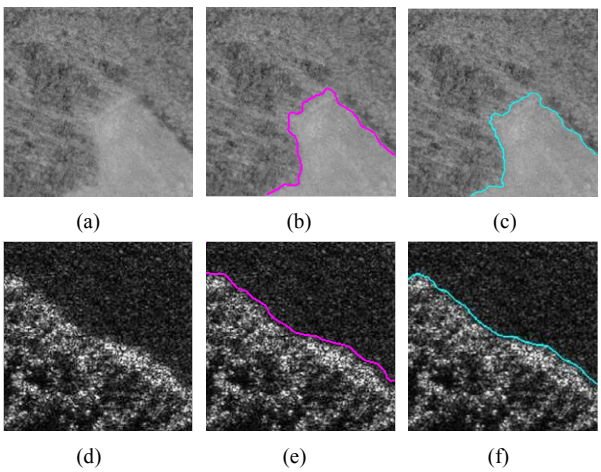
Fig.5 Segmentation results of real-world textural images: (a)(d)(g)(j)(m)(p) Original images; (b)(e)(h)(k)(n)(q) Our results; (c)(f)(i)(o) Results from Ref.[5]; (l) Result from Ref.[4]; (r) Result from Ref.[10]



(a) (b) (c) (d)

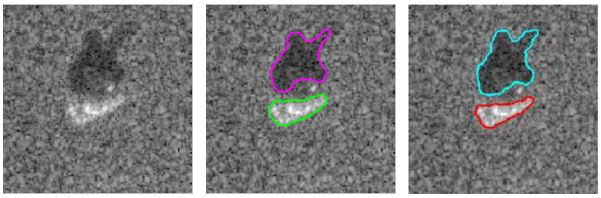
Fig.6 Segmentation results from Ref.[11]

In Figs.7 and 8, we make comparison of our method with other methods for SAR image segmentation. Fig.7(a) and (d) are two-phase SAR images. Fig.8(a) shows a real moving and stationary target acquisition and recognition (MSTAR)[22] SAR image of vehicle T72 with three regions of background, target and shadow. The accurate segmentation of such an image is quite challenging[20]. For these images, the proposed method is still effective. Our results are satisfactory and comparable to the results obtained by the method in Ref.[10].



(a) (b) (c) (d) (e) (f)

Fig.7 Two-phase SAR segmentation results: (a)(d) SAR images; (b)(e) Our results; (c)(f) Results from Ref.[10]



(a) (b) (c)

Fig.8 Three-phase SAR segmentation results: (a) An SAR image; (b) Our result; (c) Result from Ref.[10]

Identifying various textures is usually an effortless task for a human observer, but it is not an easy one in image processing and computer vision. In this paper, two novel ideas are proposed and applied to segment multiphase texture images. First, a new texture descriptor based on the patch intensity distribution information is presented. Only classification according to pixel gray level is needed in the calculation of the texture descriptor, which makes fast texture segmentation possible. The limitation of this texture descriptor lies in its invalidity when thinking of the orientation of texture. If textures differ only in orientations, our texture descriptor will fail to recognize them. This situation may occur in synthetic texture images but is rare in natural images. So the proposed texture descriptor is more suitable for the latter. This is consistent with our experimental results. Second, we reformulate the Potts model with new texture descriptor and obtain a general multiphase texture segmentation model. A fast algorithm based on split Bregman method is designed to determine a minimizer of the energy function. According to our experimental results, our method is competitive with other state-of-the-art segmentation methods for synthetic texture images and SAR images, and superior to other methods for natural images.

References

[1] S. Han, W. Tao and X. Wu, Pattern Recognition **44**, 503 (2011).

[2] L. Li, L. H. Jin, X. Y. Xu and E. M. Song, Signal Processing **93**, 2559 (2013).

[3] F. R. D. Siqueira, W. R. Schwartz and H. Pedrini, Neurocomputing **120**, 336 (2013).

[4] N. Houhou, J. P. Thiran and X. Bresson, Numerical Mathematics: Theory, Methods and Applications **2**, 445 (2009).

[5] J. Liu, X. C. Tai, H. Huang and Z. Huan, Pattern Recognition **44**, 2093 (2011).

[6] T. Ojala, M. Pietikäinen and T. T. Mäenpää, IEEE Transaction on Pattern Analysis and Machine Intelligence **24**, 971 (2002).

[7] XU Shao-ping, LIU Xiao-ping, LI Chun-quan, HU Ling-yan and YANG Xiao-hui, Journal of Optoelectronics·Laser **24**, 990 (2013). (in Chinese)

[8] WU Xiao-sheng, ZHU Shi-song, SUN Jun-ding and FAN Guo-liang, Journal of Optoelectronics·Laser **25**, 1627 (2014). (in Chinese)

[9] Y. Yang, L. Guo, T. J. Wang, W. B Tao, G. P. Shao and

Q. Feng, Image and Vision Computing **32**, 87 (2014).
[10] Fang Li and Michael K. Ng, Communications in Computational Physics **8**, 623 (2010).
[11] X. Z. Xie, J. T. Wu and M. G. Jing, Pattern Recognition Letters **34**, 1230 (2013).
[12] T. F. Chan and L. A. Vese, IEEE Transactions on Image Processing **10**, 266 (2001).
[13] L. A. Vese and T. F. Chan, International Journal of Computer Vision **50**, 271 (2002).
[14] M. Unser and M. Eden, IEEE Transactions on Systems, Man and Cybernetics **20**, 804 (1990).
[15] T. Goldstein, X. Bresson and S. Osher, Journal of Scientific Computing **45**, 272 (2010).
[16] P. Brodatz. Textures: A Photographic Album for Artists and Designers, New York: Dover Publication, 1996.
[17] D. Martin, C. Fowlkes, D. Tal and J. Malik, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition **2**, 416 (2001).
[18] J. J. Quan, X. B. Wen and X. Q. Xu, Applied Mathematics and Computation **205**, 578 (2008).
[19] G. Gao, L. Zhao, J. Zhang, D. Zhou and J. Huang, Pattern Recognition **41**, 3035 (2008).
[20] G. T. Herman and B. M. Carvalho, Simultaneous Fuzzy Segmentation of Medical Images, Handbook of Biomedical Image Analysis, Springer, 661 (2007).
[21] M. A. Savelonas, D. K. Iakovidis and D. Maroulis, Pattern Recognition Letters **29**, 1404 (2008).
[22] MSTAR Targets T72, BMP2, BTR70, SLICY, Available at: http://www.mbvlab.wpafb.af.mil/ public/ MBVDATA/, 1997 (Online).