

Image Segmentation Using A New Scalar Texture Descriptor based on Extended Structure Tensor

Shan-Qing Zhang¹, Wen-Lei Wang², Jian-Feng Lu^{1,*}, Qi-Li Zhou³, Li-Li⁴

Institute of Graphics and Image

^{1,2,3,4} Hangzhou Dianzi University of Science and Technology, Hangzhou 310018, China

* Corresponding Author: jflu@hdu.edu.cn

Received March, 2016; revised June, 2016

ABSTRACT. *The extend structure tensor (EST) is a valuable descriptor for the texture feature of image. However, EST is highly dimensional data which would bring more computational complexity and local minimum in some image. Traditional dimension reduction approaches for EST may not full use of the inherent characteristics of image. In order to solve this problem, a new scalar texture descriptor based on invariant of the EST is presented in this paper. The scalar texture descriptor is composed of the eigenvalues of EST, and used the relationship between the eigenvalues and the element of the EST to avoid direct the computation. Then the scalar texture descriptor is embedded into the segmentation model we proposed which utilizes both global and local image information. The fast energy minimization algorithm instead of the level set method is applied to solve the model. Finally, the experiments on some synthetic and real images are show the efficiency and robustness of the our model.*

Keywords: Image segmentation, Dual method, Texture descriptor, Active contour, Dual method.

1. **Introduction.** Texture image segmentation is a challenging problem in computer vision and image process areas, which usually relies on the extraction of suitable features from the image. The segmentation algorithm based on weighted sub region color histogram [1], and based on robust Gaussian mixture modeling with spatially constraints [2] has a good segmentation effect for the image with local background change, but they also have their own shortcomings, the adaptive parameters too much or can only to divided gray image, and the above segmentation algorithm for texture image segmentation results is poor, in order to solve the problem of texture feature description, Gabor filters are used to extract texture features for the segmentation [3, 4, 5] has proposed. Unfortunately, Gabor filter have the decisive drawback that they induce a lot of redundancy and feature channels, Bigun et al. [6, 7] proposed the structure tensor (ST) to discriminate textures. ST is computed by applying a Gaussian smooth function to gradient tensor. The Gaussian filters not only perform the desirable integration of multiple edge responses from complicated structure responses, such as corner, but also lead to an undesirable blurring of structure information. In [8], a nonlinear structure tensor (NLST) has been proposed, the Gaussian smooth function is replaced by the nonlinear diffusion, a vector-valued image based on the three components of nonlinear linear structure tensor (NLST) and the image intensity information embedded in the Geodesic Active Regions (GAR) models [9]. The method computation is very complicated due to consider the relations of the four channels during segmentation. In [10], only one feature based on the eigenvalue of NLST

is used to segment the complex texture image. Although the texture segmentation based on LST or NLST information has shown relatively good results, this kind of segmentation suffers from a major drawback, the LST or NLST does not include any intensity information (or color information) from the image, a significant information misuse or loss. The problem was partially solved in [11] by introducing a compact extended structure tensor (EST). Compared to the traditional structure tensor, the extended structure tensor has six characteristic channels, and the three channels contain the intensity information. The principal components analysis method (PCA) is used to reduce the dimension to the two dimensional tensor, and the segmentation speed is still very slow. In order to improve the performance, a mean texture is designed for image segmentation by using the local ChanCVese (LCV) model [12]. Once the features are extracted, a segmentation model can be formalized by the general variation framework, for example, GAR models. The GAR segmentation model, which combines the boundary and regional features, is free from the initial conditions, can automatically deal with topological changes, and can be used to deal with various frame partition problems. In this paper a novel scalar texture descriptor based on EST is presented and used as region information to segment texture images. We also propose a combinational segmentation model which utilizes both the texture information and the intensity information to improve the performance of texture segmentation. The energy functional for the proposed model consists of three terms, i.e., global term, local term and regularization term. The texture term is constructed with the GAR model, the intensity term is designed with the well-known CV model [13], and the regularization term is based on the length of evolving curve. A fast and easy implementation algorithm instead of the level set method to drive the active contour toward the minimum in our proposed energy model. The algorithm is based on the global minimum of active contour problems [14, 15, 16] and the dual formulation [17] of the Total Variation (TV) norm. The paper is organized as follows: the scalar texture descriptor based on the EST of an image is proposed in Section 2. A combinational segmentation model based on the GAR model and the CV model is addressed in Section 3. Then the fast algorithm to determine the minimum solution is developed in Section 4. Our algorithm steps are described in the section 5. Experimental results and analyses are presented in Section 6. Finally, the summary is given in Section 7.

2. Scalar texture descriptor. Structure tensor or extended structure tensor is highly dimensional data. Although it has a good texture description, but complex computation is required. In order to improve the efficiency of the image segmentation algorithm, it is necessary to reduce the dimension and the scalar quantity effectively. The original evidences show that the scalar type structure tensor and the extended structure tensor descriptors not only have a good ability to describe the texture, but also improve the efficiency of the algorithm. In this section, we give a new scalar method for the extended structure tensor, which makes full use of the invariance of tensor distance, and also improves efficiency of the algorithm.

2.1. Previous works. Given a gray image $I : \Omega \rightarrow R$, the traditional structure tensor is defined as follows:

$$J_\sigma = K_\sigma * (\nabla I \nabla I^T) = \begin{pmatrix} K_\sigma * I_x^2 & K_\sigma * I_x I_y \\ K_\sigma * I_x I_y & K_\sigma * I_y^2 \end{pmatrix}, \quad (1)$$

Where K_σ is a Gaussian kernel with standard deviation σ , ∇I is vector $\nabla I = (I_x, I_y)^T$, vector ∇I^T is transpose of vector ∇I , and subscripts denote partial derivatives. For

vector-valued image, the above definition can be extended as:

$$J = K_\sigma * \left(\sum_{i=1}^m \nabla I_i \nabla I_i^T \right), \quad (2)$$

The Gaussian smoothing may make the structure tensor suffer from the dislocation of edges, leading to inaccurate segmentation results near region boundaries. In order to overcome this drawback, the Gaussian smoothing with nonlinear diffusion can be applied. For example, the nonlinear diffusion equation may be obtained by:

$$\partial_t u_i = \text{div} \left(g \left(\sum_{k=1}^n |\nabla u_k|^2 \right) \nabla u_i \right), \quad (3)$$

Where u_i is the three components of the structure tensor. Applying (3) with initial conditions $(u_1, u_2, u_3) = (I_x^2, I_y^2, I_x I_y)$ one can obtain the nonlinear structure tensor:

$$\widehat{J} = \begin{pmatrix} \widehat{u}_1 & \widehat{u}_3 \\ \widehat{u}_3 & \widehat{u}_2 \end{pmatrix}, \quad (4)$$

Where $\widehat{}$ components denote the nonlinearly diffused components. From the above definitions, we can find that both ST and NLST have a disadvantage of not using any intensity information at all. In order to make full use of intensity information of image, the extended structure tensor (EST) was proposed in [12]. For a scalar image $I : \Omega \rightarrow R$, the EST is defined as follows:

$$J^E = K_\sigma * (vv^T) = K_\sigma * \begin{pmatrix} I_x^2 & I_x I_y & I_x \\ I_x I_y & I_y^2 & I_y I \\ I_x I & I_y I & I^2 \end{pmatrix}, \quad (5)$$

With $v = [I_x, I_y, I]^T$, Similarly, applying (3) with initial conditions $u = [I_x^2, I_y^2, I_x I_y, I_x I, I_y I, I^2]$, one can obtain a nonlinearly extended structure tensor (NLEST):

$$\widehat{J}^E = D(vv^T) = \begin{pmatrix} \widehat{I}_x^2 & \widehat{I}_x I_y & \widehat{I}_x I \\ \widehat{I}_x I_y & \widehat{I}_y^2 & \widehat{I}_y I \\ \widehat{I}_x I & \widehat{I}_y I & \widehat{I}^2 \end{pmatrix}, \quad (6)$$

The formulae (5) and (6) can be extended to vector-valued images like (2) easily and they not only contain the texture information but also the intensity information. Therefore, both EST and NLEST have stronger ability for the discrimination between different textures. However, both EST and NLEST are 3×3 matrix which has six different components. This implies that the segmentation with either EST or NLEST has to be implemented in a higher dimensional space and thus results in complex computation and multiple local minima. In order to reduce the computation complexity and the number of dimensions, we map tensor data into scalars for further segmentation. For example, a mean texture is introduced as the following:

$$T_m = \frac{1}{9} \sum_{i=1}^3 \sum_{j=1}^3 J_{ij}, \quad (7)$$

Where J_{ij} are components of the EST matrix \widehat{J}^E . The mean texture is only the average of all the components of the EST \widehat{J}^E without considering the relations of each channel of the EST \widehat{J}^E . where J_{ij} are components of the EST matrix \widehat{J}^E . The mean texture is only the average of all the components of the EST \widehat{J}^E without considering the relations of each channel of the EST \widehat{J}^E .

2.2. A novel scalar texture descriptor. Since this kind of scalarization texture approach is too rough, a more reasonable scalar textural feature has to be added. In [18], a new anisotropy measure for tensor field which is rotationally invariant has been proposed for DTI-MRI Brain data. The measure can be easily applied to the extended structure tensor. It is well known that the eigenvalues of the extended structure tensor \widehat{J}^E are tensor invariant. Given the EST matrix \widehat{J}^E , the characteristic equation as following:

$$\lambda^3 - A_1\lambda^2 + A_2\lambda - A_3 = 0 \quad (8)$$

Where

$$\begin{aligned} A_1 &= J_{11} + J_{22} + J_{33}, \\ A_2 &= J_{11}J_{22} - J_{12}J_{21} + J_{11}J_{33} - J_{13}J_{31} + J_{22}J_{33} - J_{23}J_{32}, \\ A_3 &= J_{11}(J_{22}J_{33} - J_{32}J_{23}) - J_{12}(J_{21}J_{33} - J_{31}J_{23}) + J_{13}(J_{21}J_{32} - J_{31}J_{22}). \end{aligned} \quad (9)$$

Denote the eigenvalues of J^E by $\lambda_1, \lambda_2, \lambda_3$, it is easy to prove that:

$$\begin{aligned} A_1 &= \lambda_1 + \lambda_2 + \lambda_3, \\ A_2 &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3, \\ A_3 &= \lambda_1\lambda_2\lambda_3. \end{aligned} \quad (10)$$

Instead of describing texture using $\lambda_1, \lambda_2, \lambda_3$, a new scalar texture is defined with A_1, A_2, A_3 as follows:

$$T_a = \frac{1}{6} \left[\frac{A_1 A_2}{A_3} - 3 \right], \quad (11)$$

It is easy to show that T_a is always $\lambda_{\max}/\lambda_{\min}$ and measures the magnitude of the intensity. The comparison between the mean texture T_m and the new texture T_a is given in FIGURE 1

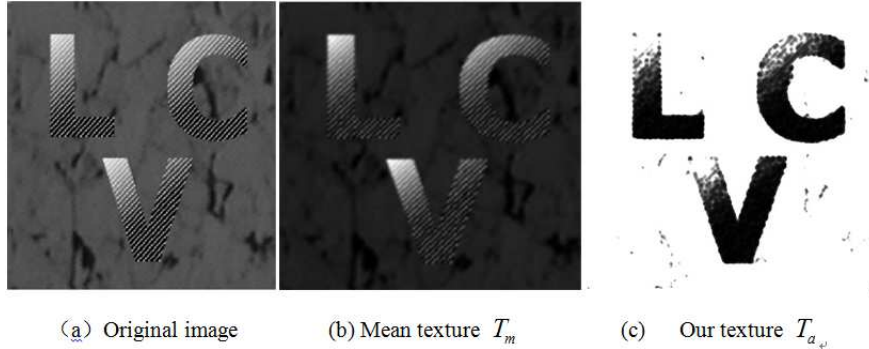


FIGURE 1. Comparison between T_m and T_a

3. A noval combined segmentation model. Once the texture feature has been selected, a segmentation model can be designed with energy minimization method. Given a gray image $I(x, y) : \Omega \rightarrow R$, let C be a closed contour in the image domain Ω , which separates Ω into two regions: $\Omega_1 = \text{inside}(C)$ and $\Omega_2 = \text{outside}(C)$. $\varphi(x)$ is the level set form of C , i.e. $C = \{x | \varphi(x) \equiv 0\}$. Embedding the texture descriptor T_a into the GAR model, we can obtain energy function:

$$E^T = - \int_{x \in \Omega_1} \log p_1(T_a(x)) dx - \int_{x \in \Omega_2} \log p_2(T_a(x)) dx, \quad (12)$$

Where p_i is the probability density function of texture descriptor T_a in Ω_i . Inserting the intensity I into the CV model, one can obtain energy function:

$$E^I = \int_{x \in \Omega_1} |I(x) - c_1|^2 dx + \int_{x \in \Omega_2} |I(x) - c_2|^2 dx, \quad (13)$$

Where c_i is the mean of image in Ω_i . In order to overcome the inhomogeneous intensity distribution in some images segmentation, Imitate [12], we mixing E^T , E^I and adding the regular term as usual, a new energy function is obtained as the following:

$$E_1 = \lambda \int_{\Omega} ds + (1 - \omega)E^T + \omega E^I, \quad (14)$$

Where λ is a constant parameter, and ω is a parameter which balances the interactions of the texture T and the intensity I . Here, we set $\omega=0.01$.

4. Fast energy minimization algorithm. Instead of using the level set method, a fast energy minimization algorithm is described, which is based on the global minimum of active contour problems [15, 16] and the dual formulation of the TV norm as proposed by Chambolle in [17]. The level set form is defined as:

$$\begin{aligned} E_2(\varphi(x)) &= \lambda \int_{x \in \Omega} |\nabla H(\varphi(x))| dx \\ &+ (1 - \omega) \left[- \int_{x \in \Omega} H(\varphi(x)) \log p_1(T_a(x)) dx - \int_{x \in \Omega} (1 - H(\varphi(x))) \log p_2(T_a(x)) dx \right] \\ &+ \omega \left[\int_{x \in \Omega} H(\varphi(x)) |I(x) - c_1|^2 dx + \int_{x \in \Omega} (1 - H(\varphi(x))) |I(x) - c_2|^2 dx \right], \end{aligned} \quad (15)$$

Where $H(\varphi(x))$ is the Heaviside function. Replacing the function $H(\varphi(x))$ with a fuzzy function $u(x) = [0, 1]$ in (15), one can obtain

$$\begin{aligned} E_3(u) &= \lambda \int_{x \in \Omega} |\nabla u(x)| dx \\ &+ (1 - \omega) \left[- \int_{x \in \Omega} u(x) \log p_1(T_a(x)) dx - \int_{x \in \Omega} (1 - u(x)) \log p_2(T_a(x)) dx \right] \\ &+ \omega \left[\int_{x \in \Omega} u(x) |I(x) - c_1|^2 dx + \int_{x \in \Omega} (1 - u(x)) |I(x) - c_2|^2 dx \right], \end{aligned} \quad (16)$$

Now, we introduce another function $v(x)$ to approximate $u(x)$, the (16) becomes

$$\begin{aligned} E_4(u, v) &= \lambda \int_{x \in \Omega} |\nabla u(x)| dx + \frac{1}{2\theta} \int_{\Omega} (u(x) - v(x))^2 dx \\ &+ (1 - \omega) \left[- \int_{x \in \Omega} v(x) \log p_1(T_a(x)) dx - \int_{x \in \Omega} (1 - v(x)) \log p_2(T_a(x)) dx \right] \\ &+ \omega \left[\int_{x \in \Omega} v(x) |I(x) - c_1|^2 dx + \int_{x \in \Omega} (1 - v(x)) |I(x) - c_2|^2 dx \right], \end{aligned} \quad (17)$$

Where $u, v \in [0, 1]$, and θ should be small enough to ensure that the solution u and v are almost equivalent. In the following, Equation (17) is solved with the Chambolle dual method. We compute u as:

$$u(x) = v(x) - \lambda \theta \operatorname{div} \mathbf{P}(x), \quad (18)$$

Where the vector $P = (P^1, P^2)$ is obtained by

$$\begin{aligned} P^0 &= 0 \\ P^{n+1} &= \frac{P^n + \tau \nabla(\text{div} P^n - v/\theta)}{1 + \tau |\nabla(\text{div} P^n - v/\theta)|}, \end{aligned} \quad (19)$$

Where τ is the time step, and div is divergence operator. Iterating v with

$$v(x) = \max \{ \min \{ u(x) - \theta r(x), 1 \}, 0 \}, \quad (20)$$

Where

$$r(x) = (1 - \omega)[- \log p_1(T_a(x)) + \log p_2(T_a(x))] + \omega [|I(x) - c_1|^2 - |I(x) - c_2|^2], \quad (21)$$

Besides, the parameters c_1, c_2, p_1, p_2 are updated as follows:

$$c_1 = \frac{\int_{\Omega} u(x) I(x) dx}{\int_{\Omega} u(x) dx}, c_2 = \frac{\int_{\Omega} (1 - u(x)) I(x) dx}{\int_{\Omega} (1 - u(x)) dx}, \quad (22)$$

$$p_1(T(x)) = \frac{\int_{\Omega} u(\hat{x}) G_{\sigma}(T(\hat{x}) - T(x)) d\hat{x}}{\int_{\Omega} u(\hat{x}) d\hat{x}}, p_2(T(x)) = \frac{\int_{\Omega} (1 - u(\hat{x})) G_{\sigma}(T(\hat{x}) - T(x)) d\hat{x}}{\int_{\Omega} (1 - u(\hat{x})) d\hat{x}}, \quad (23)$$

Where G_{σ} is a two-dimension Gaussian function.

5. Description of algorithm steps. Now, we can describe the steps of our segmentation model as follows:

Step 1 Input the original image I . If the image is texture image then compute the corresponding extended structure tensor \widehat{J}^E according to (6).

Step 2 Based on (11) compute the scalar texture descriptor T_a corresponding to extended structure tensor Extended structure tensor \widehat{J}^E and replace I with T_a .

Step 3 The obtained scalar texture descriptor is embedded into the model of the combination of our design in (14).

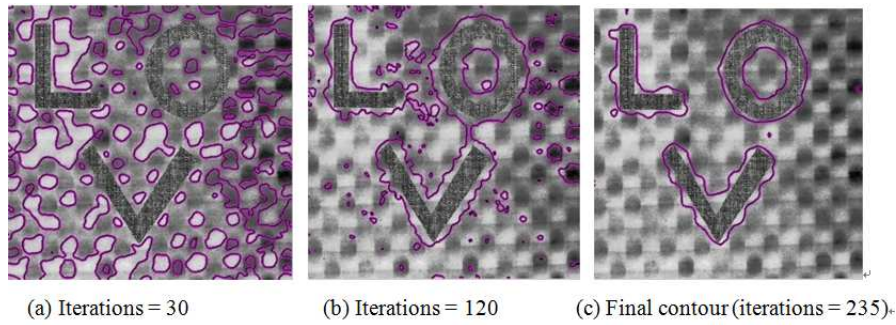
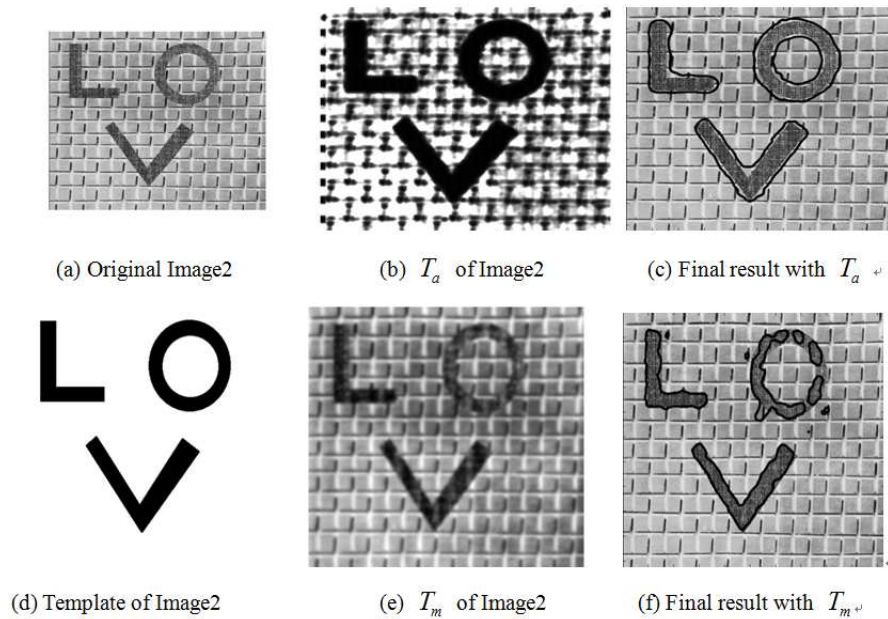
Step 4 The above model is solved by the dual method.

6. Experimental results. In this section, the experimental results of our segmentation model are presented on some synthetic and real images. The compared experiments using the synthesized texture images and the real images are carried out to reveal the powerful texture discriminating capability of our proposed model. The proposed model was implemented by Matlab 7.10.0 on a computer with Intel Core i5-2400 3.10 GHz CPU 4G RAM, and Windows 7 operating system. We used the same parameters of $\tau = 0.01, \theta = 1/120, \lambda = \frac{1}{\theta} \times 5 \times 10^{-5}$. FIGURE 2 shows the segmentation results with different iterations for a synthetic texture image by our proposed model.

The comparisons between texture descriptors T_a and T_m for several images with the new model are given from FIGURE 3 to FIGURE 5.

To further show the advantage of the texture descriptor T_m , the number of iterations, CPU time and Dice Similarity Coefficient (DSC) [19] based on T_a and T_m with our model for the images listed in Table 1. Table 1 shows that the CPU times with T_a are slightly increased at the same number of iterations while the DSCs are improved significantly.

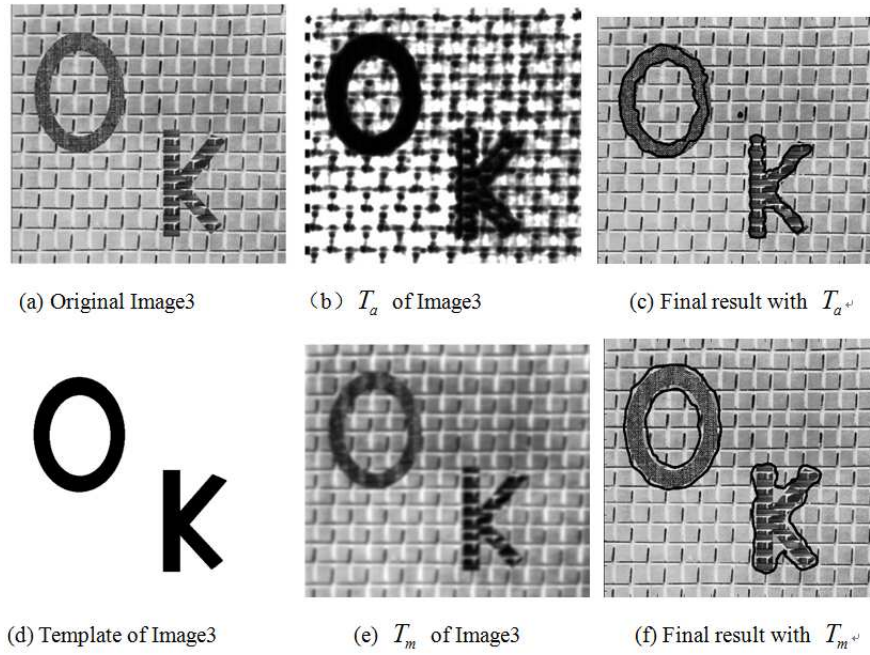
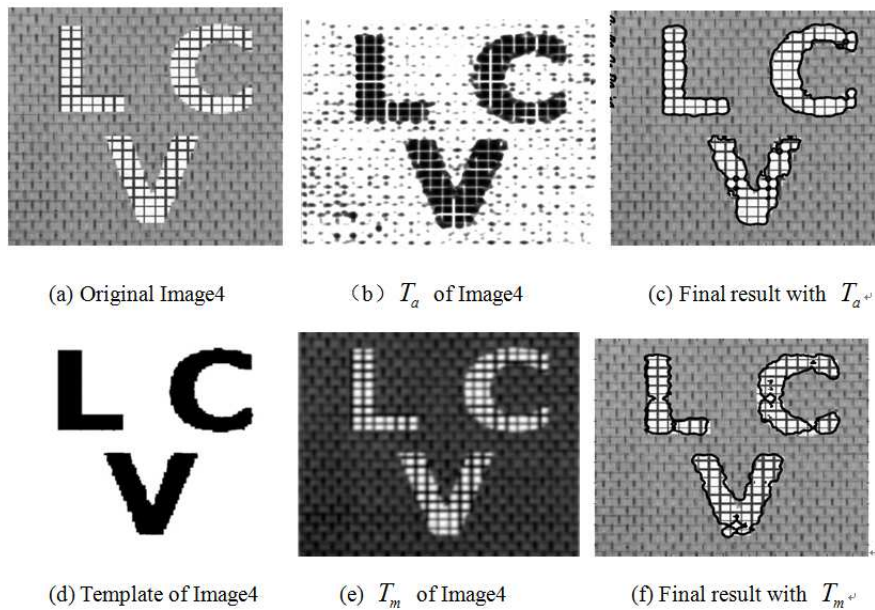
Other tests have been implemented for two synthetic scale texture images based on T_a with our model, the segmentation results are given in FIGURE 6.

FIGURE 2. Comparison between T_m and T_a FIGURE 3. Comparison between T_m and T_a TABLE 1. CPU time and DSC based on T_a and T_m with our model for the some images

Images	texture	Iterations	CPU times(s)	DSC
Image2	T_m	200	422.32	0.93
	T_a	200	429.72	0.96
Image3	T_m	200	402.96	0.92
	T_a	200	404.44	0.96
Image4	T_m	100	251.46	0.95
	T_a	100	217.79	0.97

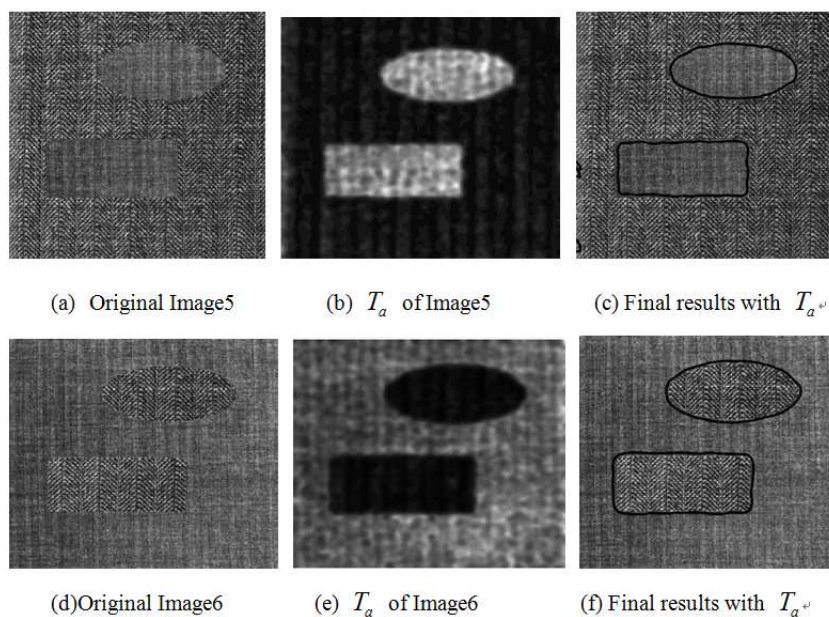
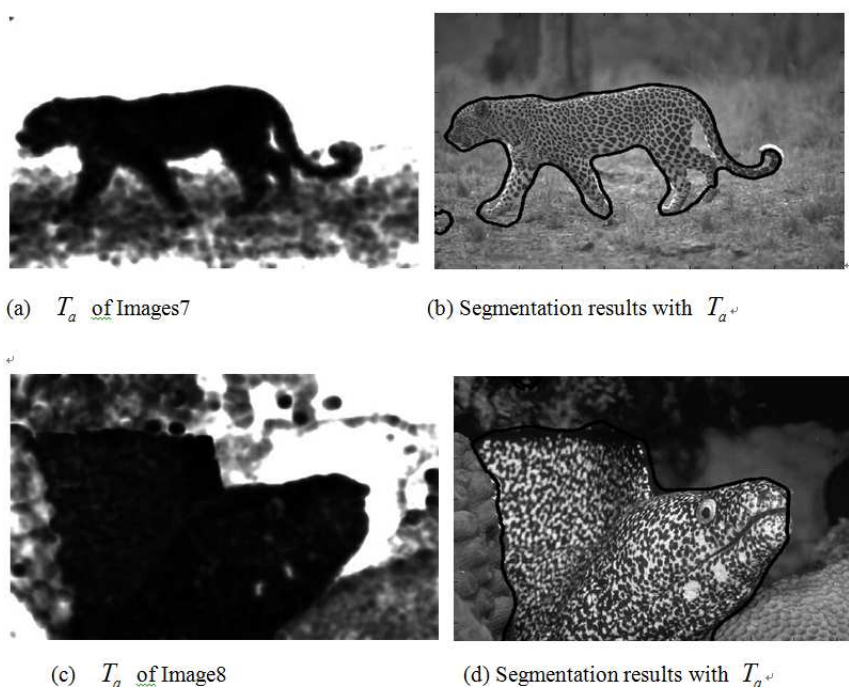
As for real texture images, some segmentation results with T_a are shown in FIGURE 7 and FIGURE 8.

7. Conclusions. In this paper, we propose a new approach which is based on the extended structure tensor to map the tensor data into a scalar without requiring computation of eigenvalues. A mixing segmentation model for texture image is introduced by applying

FIGURE 4. Comparison between T_m and T_a FIGURE 5. Comparison between T_m and T_a

the Chan-Vese(CV) model and the Geodesic Active Regions model(GAR) to the derived scalar texture characters. A fast energy minimization algorithm is then proposed with the fast dual method and used for quantitative analysis with the mean texture. Finally, the experiments on some synthetic and real images have demonstrated the efficiency and robustness of our method.

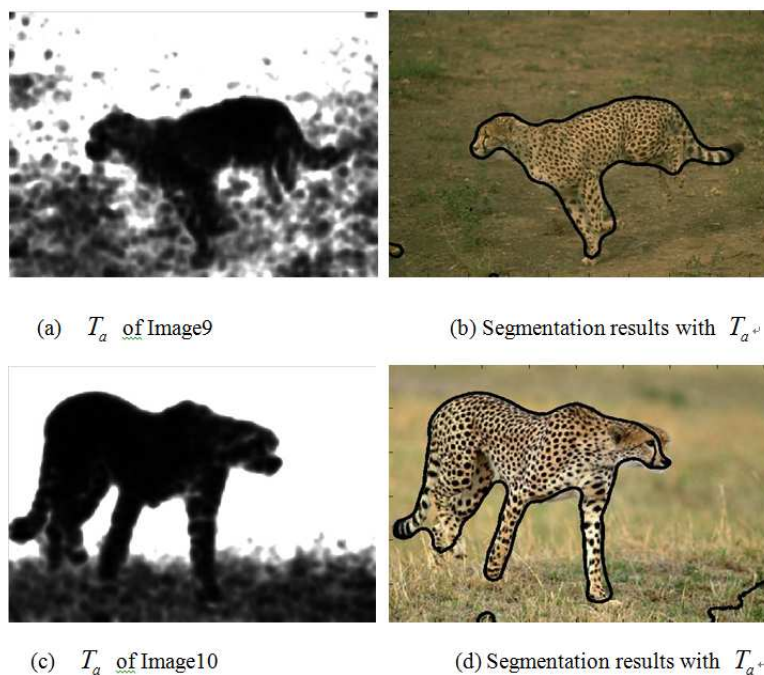
Acknowledgment. This work was mainly supported by Zhejiang Provincial Natural Science Foundation of China (No. LY14F020043) and National Natural Science Foundation

FIGURE 6. Comparison between T_m and T_a FIGURE 7. Comparison between T_m and T_a

of China (No.61370218).

REFERENCES

- [1] Y. C. Xing, E. Q. Sun, Y. Lu, Z. M. Lu, Multi-scale shot segmentation based on weighted subregion color histogram, Journal of Information Hiding and Multimedia Signal Processing, vol.6, no.3, pp.622-628, 2015.

FIGURE 8. Comparison between T_m and T_a

- [2] T. S. Xiong, Y. Y. Huang, Robust Gaussian mixture modelling based on spatially constraints for image segmentation, *Journal of Information Hiding and Multimedia Signal Processing*, vol.6, no.5, pp.857-868, 2015.
- [3] A Kaur, G Jindal, *Texture based image segmentation using Gabor filters*, *International Journal of Engineering Science and Advanced Technology [IJESAT]*, vol.3, no.3, pp.687-689, 2014.
- [4] S.W Franklin, S.E Rajan, Retinal vessel segmentation employing ANN technique by Gabor and moment invariants-based features, *Applied Soft Computing*, vol.22,no.5, pp.94-100, 2014.
- [5] K.H Zhang, H.H Song, L Zhang. Active contours driven by local image fitting energy. *Pattern recognition*, vol.43, no.4, pp.1199-1206, 2010.
- [6] J Wu, Z Feng, Z Ren, Improved structure-adaptive anisotropic filter based on a nonlinear structure tensor, *Journal of Cybernetics and Information Technologies*, vol.14, no.1, pp.112-127, 2014.
- [7] J Bign, G.h Granlund, J Wiklund, Multidimensional orientation estimation with applications to texture analysis and optical flow, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.13, no.8, pp.775-790, 1991.
- [8] M Rousson, T Brox, R Deriche. Active unsupervised texture segmentation on a diffusion based feature space, *Computer Vision and Pattern Recognition, Journal of IEEE computer society conference on*, vol.2, pp.699-704, 2003.
- [9] N Paragios, R Deriche, Geodesic active regions and level set methods for supervised texture segmentation, *International Journal of Computer Vision*, vol.46, no.3, pp.223-247, 2002.
- [10] S.Q Zhang, K.L Zhang, Texture image segmentation model based on eigenvalues of structure tensor, *Acta Electronica Sinica*, vol.41, no.7, pp.1324-1328, 2013.
- [11] R De Luisgarca, R Deriche, M Rousson, C Alberolalopez. Tensor processing for texture and colour segmentation. *Image Analysis, LNCS.3340*, pp.1117-1127, 2005.
- [12] X.F Wang, Huang D S, Xu H. An efficient local Chanese model for image segmentation *Pattern Recognition*, vol.43, no.3, pp.603-618, 2010.
- [13] S. G Liu, Y. L Peng, A local region-based ChanCvese model for image segmentation , *Pattern Recognition*, vol.45, no.7, pp.2769-2779, 2012.
- [14] X.F Wang, H Min, L Zou, Y.G Zhang, A novel level set method for image segmentation by incorporating local statistical analysis and global similarity measurement, *Pattern Recognition*, vol.48, no.1, pp.189-204, 2015.
- [15] B Conejo, S Leprince, F Ayoub, J.P Avouac, Fast global stereo matching via energy pyramid minimization, *Journal of Photogrammetric Computer Vision*, vol.2, no.3, pp.41-48, 2014.

- [16] A Beck, M Teboulle. A fast dual proximal gradient algorithm for convex minimization and applications , *Operations Research Letters*, vol.42, no.1, pp.1-6, 2014.
- [17] A Chambolle, An algorithm for total variation minimization and applications, *Journal of Mathematical imaging and vision*, vol.20, no.2, pp.89-97, 2004.
- [18] L Zhukov, K Museth, D Breen, A Barr, R Whitaker, Level set modeling and segmentation of diffusion tensor magnetic resonance imaging brain data, *Journal of Electronic Imaging*, vol.12, no.1, pp.125-133, 2003.
- [19] C.J He, Y Wang, Q Chen. Active contours driven by weighted region-scalable fitting energy based on local entropy, *Journal of Signal Processing*, vol.92, no.2, pp.587-600, 2012.