

Automatic Texture Segmentation Based on Wavelet-Domain Hidden Markov Tree

Qiang Sun, Biao Hou, and Li-cheng Jiao

Institute of Intelligent Information Processing, Xidian University,
710071 Xi'an, China
qsun@mail.xidian.edu.cn

Abstract. An automatic texture segmentation approach is presented in this paper, in which wavelet-domain hidden Markov tree (WD-HMT) model is exploited to characterize the texture features of an image, an effective cluster validity index, the ratio of the overlap degree to the separation one between different fuzzy clusters, is used to determine the true number of the textures within an image by solving the minimum of this index in terms of different number of clusters, and the possibilistic C-means (PCM) clustering is performed to extract the training sample data from different textures. In this way, unsupervised segmentation is changed into self-supervised one, and the well-known HMTseg algorithm in the WD-HMT framework is eventually used to produce the final segmentation results, consequently automatic segmentation process is completed. This new approach is applied to segment a variety of composite textured images into distinct homogeneous regions with satisfactory segmentation results demonstrated. Real-world images are also segmented to further justify our approach.

1 Introduction

Image segmentation is an important and hard problem in image analysis. Among others, texture plays an important part in low level image analysis. The image segmentation based on textural information is termed as texture segmentation, which involves the identification of non-overlapping homogeneous regions in an image.

Typically, the first step of texture segmentation is texture feature characterization, which has been discussed through various approaches by far. In this paper, wavelet-domain hidden Markov tree (WD-HMT) model is exploited to characterize texture features. The WD-HMT model [1], proposed first by Crouse *et al.* as a type of wavelet-domain statistical signal models to characterize signals through capturing the inter-scale dependencies of wavelet coefficients, has gained more and more attention from image processing and analysis communities due to its effectiveness in performing image denoising [2, 3], segmentation [4, 5, 6], texture classification [6], texture synthesis [6] and texture retrieval [7] *etc.*

Based on the WD-HMT model, one supervised image segmentation algorithm, HMTseg [4], was presented by Choi *et al.* to solve the image segmentation problem. Later, HMTseg algorithm was improved to apply to synthetic aperture radar (SAR)

image segmentation where the “truncated” HMT model [8] was proposed to reduce the effect of speckle present at fine scales.

More recently, a variety of unsupervised segmentation algorithms [9, 10, 11, 12] have been proposed one after another to extend the supervised algorithm [2] to the unsupervised one based on WD-HMT models. Zhen [9] integrated the parameter estimation and classification into one by using one multi-scale Expectation Maximization (EM) algorithm to segment SAR images on the coarse scales. In [10], Song exploited HMT-3S model [6] and the joint multi-context and multi-scale (JMCMs) approach [5] to give another unsupervised segmentation algorithm in which K-means clustering was adopted to extract the appropriate training samples for the unknown textures based on the likelihood disparity of HMT-3S model. Subsequently, Sun [11] utilized an effective soft clustering algorithm, possibilistic C-means (PCM) clustering, to further improve the unsupervised segmentation performance. Alternatively, Xu [12] has also given one unsupervised algorithm, where the dissimilarity between image blocks was measured by the Kullback-Leibler distance (KLD) between different WD-HMT models, followed by a hierarchical clustering of the image blocks at the selected scale. It should be noted that all the unsupervised segmentation algorithms above are implemented under the assumption that the number of the textures in an image is provided *a priori*, which is unpractical for automatically segmenting images in many particular application areas, such as the content-based image retrieval.

In this paper, we present an automatic texture segmentation approach based on the WD-HMT model [1]. Firstly, one global WD-HMT model is trained with the special EM algorithm in [1] with the whole image to be segmented as one texture. This model contains information from all distinct regions, and the different goodness of fit between the global model and local texture regions exists. Secondly, the true number of textures is obtained by finding the minimum of index $v_{os}(c, U)$ [13] over $c = 2, \dots, C_{\max}$ for the likelihood results of image blocks. Thirdly, PCM clustering [14] is used to extract the training sample data based on the true number of textures. Finally, WD-HMT models for different textures are re-trained with the extracted sample data, and the supervised procedures of HMTseg [4] are performed to achieve the final results with one adaptive context based fusion scheme.

The paper is organized as follows. In Section 2, WT-HMT model is briefly reviewed. Supervised Bayesian image segmentation algorithm, HMTseg, is outlined in Section 3. Automatic segmentation approach is detailed with three main procedures in Section 4. Experimental results on composite and real images are demonstrated in Section 5. Section 6 concludes this paper.

2 Wavelet-Domain Hidden Markov Tree Model

It is well known that the discrete wavelet transform (DWT) is an effective multi-scale image analysis tool due to its intrinsic multi-resolution analysis (MRA) characteristics, which can represent different singularity contents of an image at different scales and subbands. In Fig.1 (a), one quad-tree structure of wavelet coefficients is shown, which demonstrates the dependencies of wavelet coefficients at three subbands, *HL*, *LH*, and *HH*.

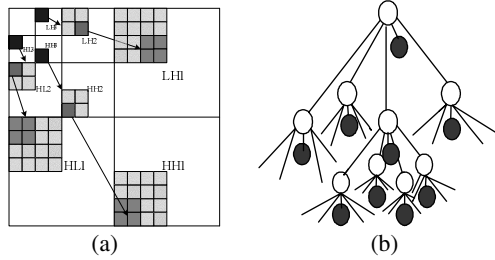


Fig. 1. (a) Quadtree structure of 2-D discrete wavelet transforms. (b) 2-D wavelet-domain hidden Markov tree model for one subband. Each wavelet coefficient (black node) is modeled as a Gaussian mixture model by a hidden state variable (white node)

For multi-scale singularity characterization, one statistical model, hidden Markov tree (HMT) model [1], was proposed to model this structure. The HMT is a multidimensional Gaussian mixture model (GMM) that applies tree-structured Markov chains across scales to capture inter-scale dependencies of wavelet coefficients [6], as shown in Fig.1 (b). In this tree-structured probabilistic model, each wavelet coefficient W is associated with a *hidden* state variable S , which decides whether it is “large” or “small”. The marginal density of each coefficient is then modeled as one two-density GMM: one large-variance Gaussian for the large state and one small-variance Gaussian for the small one. Thus, GMM can closely fit the non-Gaussian marginal statistics of wavelet coefficient.

Grouping the HMT model parameters, i.e. state probabilities for the root nodes of different quad-trees, state transition probabilities and variances for two mixed Gaussians, into one vector Θ , the HMT can be considered as one high-dimensional yet highly structured Gaussian mixture model $f(W|\Theta)$ that approximates the joint probability density function (pdf) of wavelet coefficients W . For each wavelet coefficient, the overall pdf $f(w)$ can be expressed as

$$f_W(w) = \sum_{m=1}^M p_S(m) f_{W|S}(w|S=m), \tag{1}$$

where, M is the number of states and S state variable. The model parameters in Θ are estimated by the EM algorithm in [1].

It should be noted that HMT model has one nesting structure that corresponds to multi-scale representation of an image, as shown in Fig. 2. Each subtree of the HMT is also an HMT, with the HMT subtree rooted at node i modeling the statistical characteristics of the wavelet coefficients corresponding to the dyadic square d_i in the original image.

3 Bayesian Image Segmentation Using WD-HMT

One Bayesian segmentation algorithm, HMTseg [4], was proposed to implement supervised segmentation in which the WD-HMT model [1] is exploited to characterize

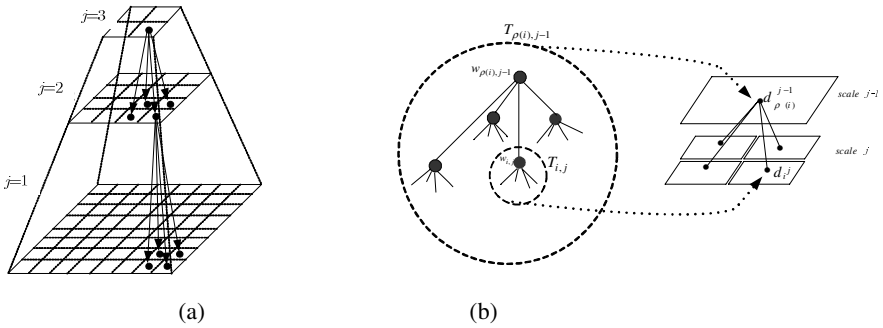


Fig. 2. Multi-scale representation of an image; (b) Correspondence of quad-tree structure of wavelet coefficients with multi-scale representation of an image

texture features and context labeling tree is built to capture the dependencies of the multi-scale class labels.

In multi-scale segmentation framework, the dyadic image squares at different scales can be obtained by recursively dividing an image into four equal sub-images. HMTseg can capture the features of these dyadic squares by the WD-HMT model. Moreover, contextual information on each dyadic square is described by one vector \mathbf{v}^j , which is derived from the labels of dyadic squares at its parent scale. Denote a dyadic square and its class label by d_i^j and c_i^j respectively, and j is the scale index.

In HMTseg [], each context vector \mathbf{v}_i^j consists of two entries, the value of the class label of the parent square and the majority vote of the class labels of the parent plus its eight neighbors.

The HMTseg algorithm relies on three separate tree structures: the wavelet transform quad-tree, the HMT, and a labeling tree [4]. As for a complete procedure, it includes three essential ingredients, i.e. HMT model training, multi-scale likelihood computation, and fusion of multi-scale maximum likelihood (ML) raw segmentations. The three main steps are summarized as follows. We refer the interested readers to Section IV in [4] to further get the knowledge on the HMTseg algorithm.

1) Train WD-HMT models for each texture using their homogeneous training images. Furthermore, Gaussian mixture is fit to the pixel values for each texture and the likelihood of each pixel is calculated to obtain the pixel-level segmentation,.

2) Calculate the likelihood of each dyadic image square d_i^j at each scale. The conditional likelihoods $f(d_i^j | c_i^j)$ for each d_i^j are obtained in this step, on which ML raw segmentation results are achieved based.

3) Fuse multi-scale likelihoods using context labeling tree to give the multi-scale maximal *a posteriori* (MAP) classification. Choose a certain suitable starting scale J such that a reliable raw segmentation can be obtained at this scale. The contextual vector \mathbf{v}^{J-1} is calculated from the class label set \mathbf{c}^J at the J -th scale. Also, the EM algorithm [4] for context labeling tree is utilized to find $p(c_i^{J-1} | \mathbf{v}_i^{J-1})$ by maximizing

the likelihood of the image given the contextual vector \mathbf{v}^{J-1} . In this step, each iteration updates the contextual posterior distribution $p(c_i | d_i, \mathbf{v}_i)$. When the process of iteration converges, determine c_i which maximizes the probability $p(c_i | d_i, \mathbf{v}_i)$. The fusion is repeated in next finer scale with the contextual vector \mathbf{v}^{J-2} computed from the label set \mathbf{c}^{J-1} at scale $J-1$. Continue the fusion process across scales until the finest scale is reached.

4 Automatic Segmentation

Automatic image segmentation using texture information means identifying all the non-overlapping homogenous regions in an image with the texture features and the number of textures unavailable. Our proposed segmentation method is made up of three steps: the determination of the number of textures utilizing v_{os} index in [13], the extraction of training sample data from different textures via the PCM clustering [14] and the supervised segmentation algorithm, HMTseg [4].

4.1 Determining the Number of Texture Categories

In this paper, the true number of textures in an image is not assumed a priori, which is different from the segmentation methods [9, 10, 11, 12], but determined using the likelihood values of image blocks at a certain suitable scale J through an effective cluster validity index for the fuzzy c -means (FCM) algorithm, v_{os} index in [13], which exploits an overlap measure and a separation measure between clusters to correctly recognize the optimal cluster number of a given data set.

Let $X = \{x_1, x_2, \dots, x_n\}$ denote a pattern set, and $x_i = [x_{i1}, x_{i2}, \dots, x_{im}]^T$ represent the m features of the i th sample. The FCM algorithm classifies the collection X of pattern data into c homogeneous groups represented as fuzzy sets $(\tilde{F}_i, i = 1, \dots, c)$. The objective of FCM is to obtain the fuzzy c -partition in terms of both the data set X and the number c of clusters by minimizing the following function

$$J_m(U, V) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - v_i\|^2, \quad \text{subject to } \sum_{i=1}^c u_{ij} = 1 \text{ for all } j. \tag{2}$$

In (2), $V = (v_1, \dots, v_c)$ is a c -tuple of prototypes, i.e. a vector of cluster centroids of the fuzzy cluster $(\tilde{F}_1, \tilde{F}_2, \dots, \tilde{F}_c)$, n is the total number of feature vectors, c is the number of classes, and $U = [u_{ij}]$ is a $c \times n$ matrix, called fuzzy partition matrix. Here, u_{ij} is the membership degree of the feature point x_j in the fuzzy cluster \tilde{F}_i and can be denoted as $\mu_{\tilde{F}_i}(x_j)$, and $m \in [1, \infty)$ is a weighting exponent, called the fuzzier, typically taken as 2.

The v_{os} index consists of two elements, an overlap measure $Overlap(c,U)$ and a separation one $Sep(c,U)$. The former measure indicates the degree of overlap between fuzzy clusters and can be obtained by calculating an inter-cluster overlap. This measure is defined as

$$Overlap(c,U) = \frac{2}{c(c-1)} \sum_{p=1}^{c-1} \sum_{q=p+1}^c \sum_{\mu \in [0.1,0.5]} \sum_{j=1}^n \delta(x_j, \mu; \tilde{F}_p, \tilde{F}_q) \times w(x_j), \quad (3)$$

where $\delta(x_j, \mu; \tilde{F}_p, \tilde{F}_q) = \begin{cases} 1 & \text{if } (\mu_{\tilde{F}_p}(x_j) \geq \mu) \text{ and } (\mu_{\tilde{F}_q}(x_j) \geq \mu) \\ 0 & \text{otherwise} \end{cases}$, and $w(x_j)$ is empirically

given a value of $0.1(\mu_{\tilde{F}_i}(x_j) \geq 0.8)$, $0.4(0.7 \leq \mu_{\tilde{F}_i}(x_j) \leq 0.8)$, $0.7(0.6 \leq \mu_{\tilde{F}_i}(x_j) \leq 0.7)$, 0 otherwise for any \tilde{F}_i . A small value of $Overlap(c,U)$ implies a well-classified fuzzy c -partition. Whereas, the latter measure $Sep(c,U)$ indicates the distance between fuzzy clusters and is defined as

$$Sep(c,U) = 1 - \min_{p \neq q} \left[\max_{x \in X} \min(\mu_{\tilde{F}_p}(x), \mu_{\tilde{F}_q}(x)) \right]. \quad (4)$$

A large value of $Sep(c,U)$ could tell one a well-separated fuzzy c -partition.

Then, the $v_{os}(c,U)$ index is expressed as the ratio of the normalized overlap measure to the separation one, i.e.

$$v_{os}(c,U) = \frac{Overlap(c,U) / \max_c Overlap(c,U)}{Sep(c,U) / \max_c Sep(c,U)}. \quad (5)$$

A small value of $v_{os}(c,U)$ indicates a partition in which the clusters are overlapped to a less degree and more separated from each other. So, the optimal value of c can be determined by minimizing $v_{os}(c,U)$ over $c = 2, \dots, C_{\max}$.

In this paper, the data set to be clustered is the likelihood values of image blocks. The true number of textures can be obtained by finding the minimum of $v_{os}(c,U)$ for the likelihood results.

4.2 Extraction of Sample Data from Different Textures

The key step for a fully unsupervised segmentation is the extraction of sample data for training different textures to obtain their HMT models used for the following supervised procedure. The input is the true number of textures in an image obtained by the cluster validity index $v_{os}(c,U)$ above. Herein, an effective soft clustering algorithm, PCM clustering [14], is exploited to extract the sample data of different textures. The objective function of the algorithm is formulated as

$$J_m(U,V) = \sum_{k=1}^N \sum_{l \in \Gamma_k} (u_{ij})^m \left\| f(y_{k,l}^{(J)} | \Theta) - f(y_k^{(J)} | \Theta) \right\|^2 + \sum_{k=1}^N \eta_i \sum_{l \in \Gamma_k} (1 - u_{ij})^m, \quad (6)$$

where U, V and m have the same meanings as those in (2), η_i is a certain positive number, and $f(y_k^{(J)}|\Theta)$ is the likelihood mean of class k at the suitable scale J , $f(y_{k,l}^{(J)}|\Theta)$ the likelihood of an image block l regarding the class k . The updated equation of u_{ij} is

$$u_{ij} = \frac{1}{1 + \left(\frac{\|f(y_{k,l}^{(J)}|\Theta) - f(y_k^{(J)}|\Theta)\|^2}{\eta_i} \right)^{\frac{1}{m-1}}}, \tag{7}$$

where η_i is defined as

$$\eta_i = \frac{\sum_{j=1}^N u_{ij}^m \|f(y_{k,l}^{(J)}|\Theta) - f(y_k^{(J)}|\Theta)\|^2}{\sum_{j=1}^N u_{ij}^m}. \tag{8}$$

PCM clustering differs from the K-means and FCM clustering algorithms since the membership of one sample in a cluster is independent of all other clusters in the algorithm. In this clustering, the resulting partition of data can be interpreted as degrees of possibility of the points belonging to the classes, i.e., the compatibilities of the points with the class prototypes [14]. Generally, more reliable and stable clustering results can be obtained with this algorithm.

The complete procedure for the PCM algorithm to implement the extraction of image sample data is listed in [14].

4.3 Adaptive Context-Based Fusion of Multi-scale Segmentation

Effective modeling of contexts for each dyadic square d_i is crucial to effectively fuse the raw segmentations from coarse scale to fine one to obtain a satisfactory result in the multi-scale fusion step. In the original HMTseg method [4], the context v_i^j is specified as a vector of two entries consisting of the value of class label $C_{\rho(i)}$ of the parent square and the majority vote of the class labels of the parent plus its eight neighbors, as illustrated in Fig.3 (a). This simplified context is typically effective for images consisting of separate large homogeneous textures since it focuses on the information of class labels at coarse scales. However, the segmentation results might be unsatisfactory when complicated structures occur in an image, such as most real-world images. In [5], Fan proposed a joint multi-context and multi-scale (JMCMs) approach to Bayesian image segmentation using WD-HMT models, where three contexts (context-2, context-3 and context-5 shown in Fig. 3) are exploited sequentially to fuse the raw segmentation results across scale. However, the computation cost is too expensive, which renders this approach unpractical in real-time image segmentation applications. Herein, one adaptive context model, as shown in Fig. 3(d), is given to fully incorporate both the information of the class labels at the coarse scale and the

information at the fine scale to further improve the segmentation performance. In this context model, the context vector for each image block contains two entries of which the first element $V1$ is defined in the same way with [4], whereas the other one $V2$ is obtained by the compromise between the coarse scale and fine scale. Generally speaking, if the dominant label Ω_1 at the coarse scale is identical with Ω_2 at the fine scale, $V2$ is established like context-2; otherwise, $V2$ is assigned Ω_2 . In this way, the new context could adaptively make a trade-off between the parent-scale ML classification results and those at the child scale. It is expected that better segmentation results could be achieved.

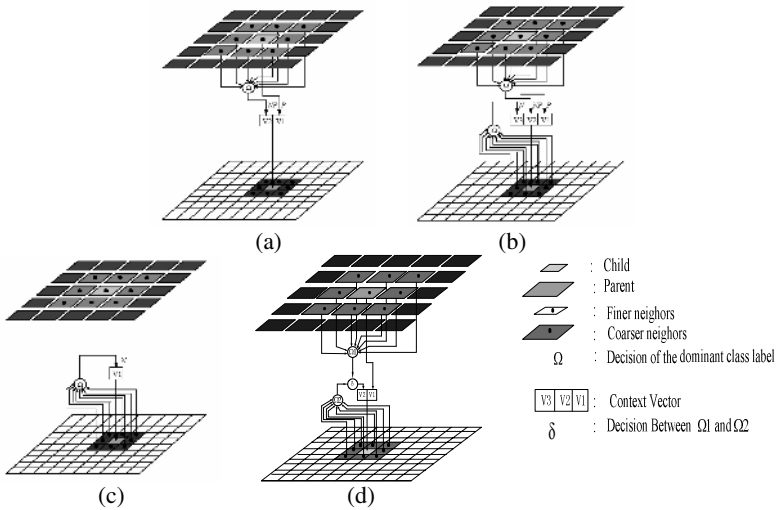


Fig. 3. Context models for inter-scale raw segmentation fusion. (a) Context-2 in [2]; (b) Context-3 in [11]; (c) Context-5 in [11]; (d) Context proposed.

5 Experimental Results

We testified our approach on composite texture images with the size of 256×256 pixels, which are made up of the original textures from Brodatz album [15]. Here, four composite textured images, consisting of 2, 3, 4 and 5 classes of homogeneous textures respectively, are shown in Fig. 4.

Originally, all the textured images are decomposed into four levels by discrete wavelet transform (DWT). The true number of the textures is determined by the disparity of the likelihoods for different image blocks using the cluster validity index $v_{os}(c, U)$ at the suitable, $J = 4$ here, which is the coarsest scale. The number of cluster goes through from 2 to 10 (C_{max}), and the optimal (true) number of the textures in an image is found by evaluating the minimum of $v_{os}(c, U)$. Then, the PCM clustering of the model likelihoods is conducted at the scale J .

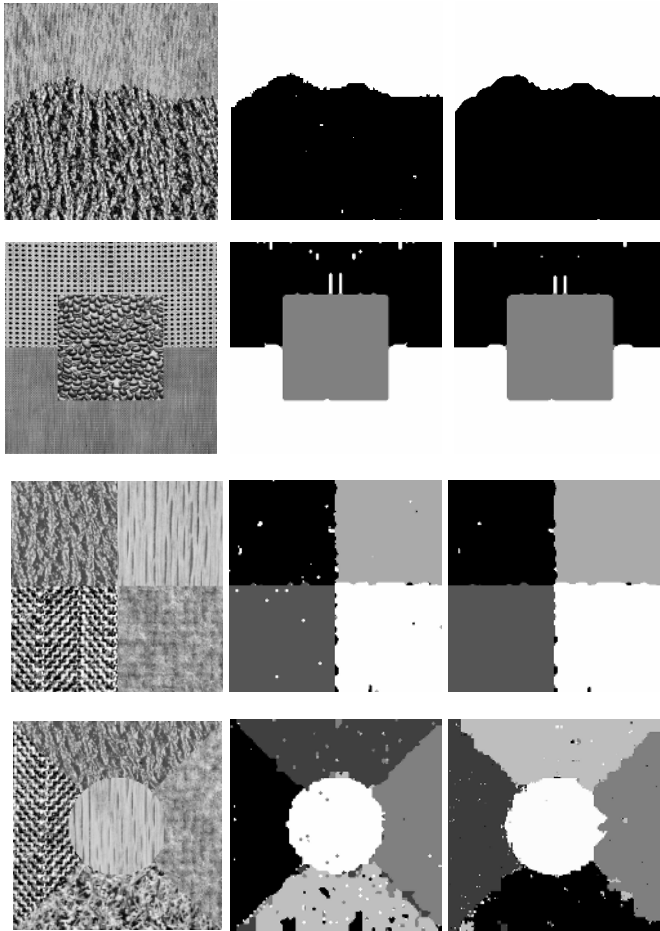


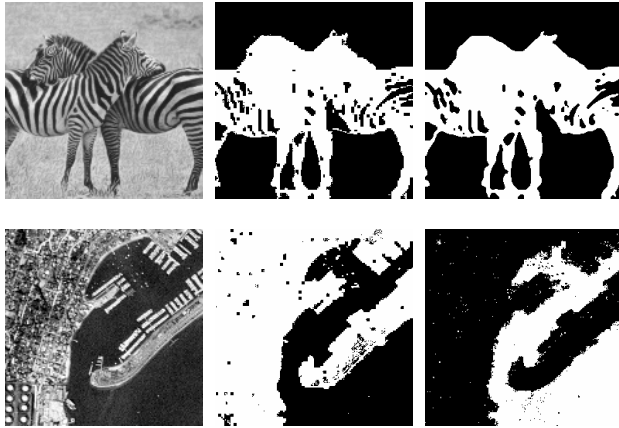
Fig. 4. Four composite texture images and their segmentation results with the proposed approach (the second column) and supervised HMTseg algorithm [4] (the third row)

In Table 1, the values of $v_{os}(c,U)$ in terms of different c for the four composite textured images in Fig. 4 are tabulated, of which the minimum of $v_{os}(c,U)$ is marked with boldface. It can be seen that all the true number of textures in these images have been correctly determined. Moreover, we also applied our method to other composite textures with a return of over 70% correctly detected number of textures obtained.

Fig. 4 also demonstrates the final segmentation results for the four composite textures with the proposed approach and the supervised HMTseg algorithm in [4]. The results demonstrate that the segmentation performance of our approach is basically satisfactory and favorably compares with the results with HMTseg. The rate of misclassified pixels for the four images is given in Table 1. Our approach gives the error percentage of below 8% for all tested composite textured images, which is basically feasible for practical applications. Meanwhile, the segmentation results for real-world images are shown in Fig. 5 with similar performances with the HMTseg algorithm [4].

Table 1. Values of $v_{os}(c, U)$ in terms of different c for the four composite textured images in Fig. 4 and their rate of misclassified pixels

Image	Number of textures	Values of $v_{os}(c, U)$ for $c=2, \dots, 10$	Rate of misclassified pixels
Composite-2	2	0.0931(2) , 0.1257(3), 0.1969(4), 0.3150(5), 0.2500(6), 0.3950(7), 0.6881(8), 1.2424(9), 1.0168(10)	0.59%
Composite-3	3	1.9374(2), 0.0308(3) , 0.3746(4), 0.2774(5), 0.1936(6), 0.1184(7), 0.0908(8), 0.0659(9), 0.0586(10)	4.65%
Composite-4	4	1.7898(2), 1.0708(3), 0.0623(4) , 0.1634(5), 0.1089(6), 0.1226(7), 0.1049(8), 0.0770(9), 0.0996(10)	3.52%
Composite-5	5	1.9502(2), 0.0460(3), 0.0131(4), 0.0111(5) , 0.0642(6), 0.0353(7), 0.0231(8), 0.0659(9), 0.0868(10)	6.79%

**Fig. 5.** Real-world images (Zebra and aerial-photo images) and their segmentation results with the proposed approach (the second column) and the HMTseg algorithm [4] (the third row)

6 Conclusions

In this paper, an automatic texture segmentation is developed by characterizing the texture features using WD-HMT model, determining the number of textures with the cluster validity index v_{os} , and extracting the sample data from different textures by means of PCM clustering. Experimental results demonstrated that the proposed method can detect correctly the number of textures and provide good segmentation results on textured images. The further work is concerned with the use of more accurate statistical model describing texture feature, such as HMT-3S model [6].

References

1. Crouse, M.S., Nowak, R.D., Baraniuk, R.G.: Wavelet-Based Signal Processing Using Hidden Markov Models. *IEEE Trans. on Signal Processing*. 46 (1998) 886–902
2. Romberg J.K., Choi, H., Baraniuk, R.G.: Bayesian Tree-structured Image Modeling Using Wavelet-Domain Hidden Markov Models. *IEEE Trans. on Image Processing*. 10 (2001) 1056–1068
3. Fan, G.L., Xia, X.G.: Image Denoising Using Local Contextual Hidden Markov Model in the Wavelet Domain. *IEEE Signal Processing Letters*. 8 (2001) 125–128
4. Choi, H., Baraniuk, R.G.: Multi-scale Image Segmentation Using Wavelet-Domain Hidden Markov Models. *IEEE Trans. on Image Processing*. 10 (2001) 1309–1321
5. Fan, G.L., Xia, X.G.: A Joint Multi-Context and Multi-Scale Approach to Bayesian Image Segmentation. *IEEE Trans. on Geoscience and Remote Sensing*. 39 (2001) 2680–2688
6. Fan, G.L., Xia, X.G.: Wavelet-Based Texture Analysis and Synthesis Using Hidden Markov Models. *IEEE Trans. on Circuits and Systems*. 50 (2003) 106–120
7. Do, M.N., Vetterli, M.: Rotation Invariant Texture Characterization and Retrieval Using Steerable Wavelet-Domain Hidden Markov Models. *IEEE Trans. on Multimedia*. 4 (2002) 517–527
8. Venkatachalam, V. , Choi, H. , Baraniuk, R.G.: Multi-scale SAR Image Segmentation Using Wavelet-Domain Hidden Markov Tree Models. In *Proc. of SPIE*, 4053 (2000) 1605–1611
9. Zhen, Y., Lu, C.C.: Wavelet-Based Unsupervised SAR Image Segmentation Using Hidden Markov Tree Models. In *Proc. of International Conference on Pattern Recognition*. 2(2002) 729–732
10. Song, X.M., Fan, G.L.: Unsupervised Bayesian Image Segmentation Using Wavelet-Domain Hidden Markov Models. In *Proc. of International Conference on Image Processing*. 2 (2003) 423–426
11. Sun, Q., Gou, S.P., Jiao, L.C.: A New Approach to Unsupervised Image Segmentation Based on Wavelet-Domain Hidden Markov Tree Models. In *Proc. of International Conference on Image Analysis and Recognition*. 3211 (2004) 41–48
12. Xu, Q., Yang, J., Ding, S.Y.: Unsupervised Multi-scale Image Segmentation Using Wavelet Domain Hidden Markov Tree. In *Proc. of the 8th Pacific Rim International Conferences on Artificial Intelligence*. 3157 (2004) 797–804
13. Kim, D.W., Lee, K.H., Lee, D.: On Cluster Validity Index for Estimation of the Optimal Number of Fuzzy Clusters. *Pattern Recognition*. 37 (2004) 2009–2025
14. Krishnapuram, R., Killer, J.M.: A Possibilistic Approach to Clustering. *IEEE Trans. on Fuzzy System*. 1 (1993) 98–110
15. Brodatz, P.: *Textures: A Photographic Album for Artists & Designers*. Dover Publications, Inc., New York, 1966