

Object-Based and Semantic Image Segmentation Using MRF

Feng Li

Shanghai Zhongke Mobile Communication Research Center, Shanghai Division, Institute of Computing Technology, Chinese Academy of Sciences, Shanghai 201203, China

Institute for Pattern Recognition & Artificial Intelligence, State Education Commission Laboratory for Image Processing & Intelligence Control, Huazhong University of Science and Technology, Wuhan 430074, China
Email: life1972@hotmail.com

Jiaxiong Peng

Institute for Pattern Recognition & Artificial Intelligence, State Education Commission Laboratory for Image Processing & Intelligence Control, Huazhong University of Science and Technology, Wuhan 430074, China
Email: jiaxpeng@sohu.com

Xiaojun Zheng

Shanghai Zhongke Mobile Communication Research Center, Shanghai Division, Institute of Computing Technology, Chinese Academy of Sciences, Shanghai 201203, China
Email: frank.zheng@cmcr.cn

Received 6 December 2002; Revised 3 September 2003

The problem that the Markov random field (MRF) model captures the structural as well as the stochastic textures for remote sensing image segmentation is considered. As the one-point clique, namely, the external field, reflects the priori knowledge of the relative likelihood of the different region types which is often unknown, one would like to consider only two-pairwise clique in the texture. To this end, the MRF model cannot satisfactorily capture the structural component of the texture. In order to capture the structural texture, in this paper, a reference image is used as the external field. This reference image is obtained by Wold model decomposition which produces a purely random texture image and structural texture image from the original image. The structural component depicts the periodicity and directionality characteristics of the texture, while the former describes the stochastic. Furthermore, in order to achieve a good result of segmentation, such as improving smoothness of the texture edge, the proportion between the external and internal fields should be estimated by regarding it as a parameter of the MRF model. Due to periodicity of the structural texture, a useful by-product is that some long-range interaction is also taken into account. In addition, in order to reduce computation, a modified version of parameter estimation method is presented. Experimental results on remote sensing image demonstrating the performance of the algorithm are presented.

Keywords and phrases: semantic and structural segmentation, MRF, Wold model, remote sensing image.

1. INTRODUCTION

In this paper, remote sensing image segmentation based on the Markov random field (MRF) is considered. Many approaches have used MRF as a label process (as discussed in [1, 2, 3, 4, 5, 6, 7, 8, 9]), including the application to extract urban areas in remote sensing images (as discussed elsewhere in [5, 10, 11]). This is because exploiting MRF offers several advantages over simple segmentation algorithms. First, the segmentation for the object in a remote sensing image depends not only on the gray level, but also on other features such as texture, which can be viewed as realizations

from a parametric probability distribution model in the image space. Second, this approach is flexible because it has a few number of parameters to set. Finite number of parameters characterizing spatial interactions of pixels is used to describe an image region. Also, the constraint of smoothness is meant to express the implicit assumption for texture segmentation, that is, each separated region has to extend over a significant area. Isolate labels and very small regions are disallowed because the texture pattern essentially can be discerned only in a large enough area. There are two basic methods for the usage of the MRF model. First (as discussed elsewhere in [12, 13]), parameters are extracted as

texture features, including mean, variance, potential parameters combined with other features, and then clustering criteria are employed to classify the image. Its advantage is simple computation. Another method (as discussed elsewhere in [1, 2, 3, 4, 5, 6, 7, 8, 9]) uses double random fields based on Bayesian framework. The advantage of this method is that prior information can be easily incorporated. Some high-level prior information can be incorporated into this framework, but the computation for combination optimization is undesirable.

There are some deficiencies of the MRF model in image analysis. Firstly, the hypothesis of homogeneous property for random field does not accord with most practical images, leading to smoothness in the texture edge. However, if a nonhomogeneous random field is used, which is relative to position and orientation, there are a number of parameters which inevitably brings about enormous computation. Secondly, Markovian prior model is a low-level prior model. It is short of semantic information and will lead to a condition in which the segmented regions are often not consistent with the object. Thirdly, as the single-clique potential prior information, namely, the external field, is often unknown, the use of only pairwise interaction in the Markovian model will lead to a result in which it cannot accurately capture the structural component of the texture.

These questions cause poor quality segmentation or increase the computation time. As the segmentation process is a basic step followed by other image analyses such as compression and interpretation, an improved method is needed. It is well known that a difficulty in using MRF model, including single-clique potential, is the introduction of appropriate prior information of single-pixel cliques. As discussed by Picard [14], the authors conclude that if it were not for competition from the internal field, the synthesized random field would align itself perfectly with the desired external field. They suggest that nonhomogeneous external field can be set to the value in some reference image. But they did not give such a reference image; in addition, they cannot estimate the relative strengths of the two fields. In this study, we address and settle several issues left open there. In addition, we apply this idea to image segmentation. We will adopt a kind of Wold decomposition which can obtain pure random field and structural field. The main contribution of this paper is to extract structural component as a reference image of the external field of the MRF model. We thus incorporate the structural component to the segmented image.

Most natural textures can be modeled as a superposition of two independent random fields (as discussed by Francos et al. in [15]): a spatially homogeneous field and a spatial singularity component. The spatial singularity field includes the local structural components of the texture, which preserve the perceptual property, such as periodicity, directionality, and randomness. By using the decomposition, the stochastic component can be captured while the structural texture is also described. Following this, we can model different components of the texture. As discussed by Francos et al.

in [16], it was shown that the decomposition fits not only the homogeneous random field, but also the nonhomogeneous random field. Contrary to space domain MRF model, Wold model is a frequency domain model and it has a global characteristic such as periodicity.

Many researchers study this model for segmentation and classification (as discussed elsewhere in [12, 13, 17]). Lu in [12] extracts Wold feature for unsupervised texture segmentation, but he adapts the clustering method by combining Wold feature with wavelet features and MRSAR parameter features. In [13], different types of image features are aggregated for classification by using a Bayesian probabilistic approach. In [17], rotation and scaling invariant parameters are used. A tested texture image can be correctly classified even if it is rotated and scaled. In this paper, we will incorporate Wold decomposition into Bayesian framework as structural prior information.

The paper is organized as follows. In Section 2, we look back to the MRF-based double random fields segmentation method. In Section 3, we describe how to capture a structural texture based on the Bayesian framework. Wold decomposition is presented in Section 4. Section 5 is devoted to a modified method to estimate the model parameters. In Section 6, segmentation results are reported for remote sensing image. These results are compared with the performance of the existing algorithm. Finally, in Section 7, we conclude our presentation with remarks on this work.

2. MARKOV RANDOM FIELD

2.1. Label field model

We use the MRF to model the label field X . The conditional distribution of a point, given all other points in the field, is only dependent on its neighbors. That is, $P(x_s | x_{L-s}) = P(x_s | x_{N_s})$ for all $s \in L$. A clique c is a subset of points in L such that if s and r are two points in c , then s and r are neighbors. Notice that the set of all cliques is induced by the neighborhood system. According to the Hammersley-Clifford theorem, for a given neighborhood system, $P(x)$ can be expressed by Gibbs distribution in the form

$$P(x) = \frac{1}{z} \exp \left\{ -\frac{1}{T} \sum_{c \in C} V_c(x_c) \right\}, \quad (1)$$

where the function V_c is an arbitrary function of the values of x on the clique c , and z is a normalizing constant. The constant T is physically analogous to temperature, and the exponential $U(x) = \sum_{c \in C} V_c(x_c)$ is physically analogous to energy. C is defined as the set of all cliques associated to L , and the summation is taken over all cliques C . A relatively simple type of discrete-valued MRF, called multilevel logistic (MLL) field, is found to be appropriate for modeling region formation in image segmentation. For our application, the only nonzero potentials of the MLL are assured to be those that correspond to one-and two-pixel cliques. These cliques belong to the second-order neighborhood system.

2.2. Texture and noise model

Given a known label realization x , we assume that the observed image y is a realization of the random field Y defined on lattice L . A conventional AR model is described as

$$y(s) = \sum_{r \in \{(i,j)\}} a_{k,r}(s) y(s-r) + w_k(s). \quad (2)$$

For residual image process, we have

$$y(s) - \mu_k(s) = \sum_{r \in \{(i,j)\}} a_{k,r}(s) (y(s-r) - \mu_k(s)) + w_k(s), \quad (3)$$

where r is the offset of s , $w_k(s)$ is a white Gaussian noise with zero mean and variance $\sigma_k(s)$, and its matrix form is

$$A(g - \mu) = w - A_0(y_0 - \mu(0)). \quad (4)$$

Nonzero elements in the matrices A and A_0 come from $a_k(s)$ in (5) and $y_0 - \mu(0)$ is the boundary condition on lattice L . Since A_0 is usually not a square matrix, we cannot replace the likelihood function by w . But we can neglect $y_0 - \mu(0)$ assuming that L is very large or periodic and then $w = A(y - \mu)$. From (5), matrix A is a lower triangular matrix and its diagonal entries are 1's, so A is always nonsingular. The conditional distribution of y is

$$\begin{aligned} P(y|x) &= |A|^{-1} P(w) \\ &= \prod_s \frac{1}{\sqrt{2\pi\sigma_k^2(s)}} \exp \left[-\frac{1}{2} \sum_s \frac{w_k(s)^2}{\sigma_k^2(s)} \right]. \end{aligned} \quad (5)$$

This results in conditional log likelihood

$$\log P(y|x) = -\frac{1}{2} \sum_s \left\{ \frac{w_k^2(s)}{\sigma_k^2(s)} + \log(\sigma_k^2(s)) + \log(2\pi) \right\}. \quad (6)$$

The above formulas show a Gaussian causal AR model with nonstationary mean and nonstationary variance. The parameter set used at the point $s \in L$ is $\theta_{y|x}(s)$. Each parameter vector $\theta_{y|x}(s)$ contains the mean $\mu_k(s)$, the variance $\sigma_k(s)$, and the prediction coefficients $a_{k,r}$.

3. STRUCTURAL SEGMENTATION

The MRF model with only pairwise clique potential cannot capture particular direction as well as periodicity. When this model is applied to the structural pattern, the resulting synthesized patterns are not visually similar to the original. In addition, the usage of only pairwise statistics in the model leads to smoothness at the edge of texture. In order to solve this problem, a single-clique potential should be considered in the model. As prior information of the percentage of each region is unknown, in [14], Picard introduces the concept of reference image. Furthermore, we set the nonhomogeneous external field to the values in some reference image y_r and consider the internal field as homogeneous. Hence the external field $\alpha_s = y_{cs}$, the gray-level value at site s in the image y .

According to the Bayesian framework, we have

$$\begin{aligned} p(x|y) &\propto p(y|x)p(x), \\ p(x) &= \frac{1}{Z} \exp \left(-\frac{\nabla E(x)}{T} \right), \\ p(x|y, \alpha_k) &\propto p(y|x, \alpha_k)p(x), \end{aligned} \quad (7)$$

where $\nabla E = \nabla E_1 + \nabla E_2$;

$$\begin{aligned} P(y|x) &= |A|^{-1} P(w) = \prod_s \frac{1}{\sqrt{2\pi\sigma_k^2(s)}} \exp \left[-\frac{1}{2} \sum_s \frac{w_k(s)^2}{\sigma_k^2(s)} \right], \\ E_1(x) &= -\frac{1}{2} \left[\sum_s \frac{w_k(s)^2}{\sigma_k^2(s)} \right]^2, \end{aligned} \quad (8)$$

where

$$\begin{aligned} w_k(s) &= y_s - \mu_k(s) + \sum_{r>0} a_{k,r} (y_{s-r} - \mu_k(s)), \\ E_2(x) &= -\sum_{s \in S} \left(\alpha_k x_s + \sum_{r \in N_s} \beta_s \delta(x_s, x_r) \right) \\ &= -\sum_{s \in S} \left(\gamma y_s x_s + \sum_{r \in N_s} \beta_s \delta(x_s, x_r) \right), \end{aligned} \quad (9)$$

where γ is the proportion between the external and internal fields, β_s is the nonnegative parameter of MRF, and α is the external field. Although one can synthesize a sample from any energy range of the Gibbs distribution, the most probable samples correspond to those with the least energy. The internal field product term $\sum_{r \in N_s} \beta_s \delta(x_s, x_r)$ has been shown to be maximized when the texture in the image forms configuration which maximizes its disperse so that the minimum energy internal field will have minimal length boundaries between pairs of texture. The product is maximized when the same texture is most likely to form. The internal field product term $\alpha y_s x_s$ is the contrary; if not for the competition from the internal field product, the synthesized random field would align itself perfectly with the desired external field. It shows that the internal field describes the structural texture and it is important. In Section 4, the internal field will be obtained by Wold model decomposition.

Our segmentation is essentially based on the texture structure. However, since we are only interested in finding urban areas, we consider the problem of urban area detection as a scene-labeling problem, where each pixel in the image is assigned a label indicating which class the urban areas and the nonurban areas belong to. The results are visually quite similar to the actual texture classification and somewhat semantic for identifying properties of urban areas. So we refer to our method as object-based and semantic image segmentation.

Structural information, associated with common sense knowledge, can be helpful to obtain a coherent interpretation of the whole scene. The geometrical shape of urban areas is better preserved. For such image, we can identify classes

of data-type and classes of semantics. Classes like texture or smooth are data-type classes and classes like agricultural, urban are semantics classes. The classes of semantics are often associated with a specific data-type class.

4. WOLD MODEL

Remote sensing image can be regarded as texture, including structural or stochastic texture. Because many textures include the two components simultaneously, Francos [15] presents a new model: Wold model which can capture random, directional, and periodical textures, and can preserve the perceptual property of the image. Let $y(n, m)$ be a realization of real-valued, regular, and homogeneous random field and $F(\omega, \nu)$ a spectral distribution function. It can, respectively, uniquely be decomposed as

$$y(n, m) = w(n, m) + h(n, m) + e(n, m), \quad (10)$$

where w is a purely random field, while the structural random field includes h and e . h is a half-plane structural random field, which is represented by harmonic field, and e is called the generalized evanescent field:

$$\begin{aligned} h(n, m) &= \sum_{k=1}^p \{C_k \cos 2\pi(n\omega_k + mv_k) + D_k \sin 2\pi(n\omega_k + mv_k)\}, \\ e(n, m) &= s(n) \sum_i \{A_i \cos 2\pi mv_i + B_i \sin 2\pi mv_i\}, \end{aligned} \quad (11)$$

where C_k, D_k are mutually orthogonal random variables; A_i, B_i are mutually orthogonal random variables; and $s(n)$ is a purely 1D random process.

Starting from the original image, Gaussian taper is applied to reduce the edge effect. The theorem described above is then used to decompose the original image. When the decomposition is finished, we proceed to extract the harmonic and directional features in the structural random field by employing maximum spectral peak and Hough transformation, respectively. Francos presented an algorithm to estimate parameters of the structural field, which describe the structural texture employing the maximum likelihood (ML) estimation method. A simplified method can be used here to approximate the parameter.

The value of (w_k, v_k) can be obtained by solving the following equation:

$$(w_k, v_k) = \arg \max_{(w, v)} | \text{DFT}(y(n, m)) |^2. \quad (12)$$

In iteration, the frequency of the dominant harmonic component is estimated by

$$\begin{aligned} C_k &= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} y(n, m) \cos(w_k, v_k), \\ D_k &= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} y(n, m) \sin(w_k, v_k), \end{aligned} \quad (13)$$

where N, M are the sizes of the image. Let A be the sum matrix in Hough transformation:

$$(\rho_i, \theta_i) = \arg \max_{(\rho, \theta)} A. \quad (14)$$

(w_i, v_i) can be obtained by inverse transformation:

$$\begin{aligned} \rho_i &= w_i \cos \theta_i + v_i \sin \theta_i, \\ (w_i, v_i) &= \arg \max_{(w, v) - (w_k, v_k)} | \text{DFT}(y(n, m)) |^2. \end{aligned} \quad (15)$$

5. PARAMETER ESTIMATION

Least square parameter is estimated as follows (as discussed by Kashyap and Chellappa in [18]):

$$\theta^* = \left[\sum_{\Omega} Q(n, m) Q^T(n, m) \right]^{-1} \left(\sum_{\Omega} Q(n, m) y(n, m) \right), \quad (16)$$

where $Q(n, m) = [y(n+1, m) + y(n-1, m); y(n, m+1) + y(n, m-1)]$ and Ω represents all the pixels in the image. This method is simple to calculate, but it is not consistent. So, we will employ the ML estimation method. Because remote sensing image is large and complex as in Figure 1, MPL estimation converges to the true value with probability 1.

Because the parameter estimation scheme will take undesirable calculation time, a faster version of parameter estimation method is needed. In this paper, we use a modified simultaneous parameter estimation and segmentation. The parameter set used in formulas (8) and (9) is $\theta = (\theta_x, \theta_{y|x})$, where the parameter vector $\theta_{y|x}$ contains the mean μ_k , the variance σ , and the prediction coefficients $a_{k,r}$; the parameter vector θ_x contains the parameter β_s of MRF and γ .

As simulated annealing (SA) takes a long time to converge to the maximum of $\Pi_{s \in S} p(x_s | x_{N_s})$ over the parameter vector θ , we employ ICM-SA method, that is, initial values for the parameters are computed by performing ICM, then SA is implemented. Because ICM cannot perform backtracking, the initial condition is crucial. In [7], Pappas presents an adaptive segmentation method. There, initial parameters are presented as follows: according to the four-color theorem, the texture class number $K = 4$ is a suitable choice. Strictly speaking, the number of classes K should also be considered as an unknown parameter which has to be estimated from the image. In general, one can minimize the AIC information criteria to find the number of classes K (as discussed by Zhang et al. in [9]). The variance $\sigma = 7$, and the label field model parameter $\beta_s = 0.5$ for every s . Increasing σ^2 is equivalent to increasing β_s . The author considers these parameters as robust for most images. We adopt these values above to achieve a good initial segmentation and reduce the iteration number.

In order to achieve the desired maximization, we use the metropolis algorithm to implement ICM-SA (as discussed by

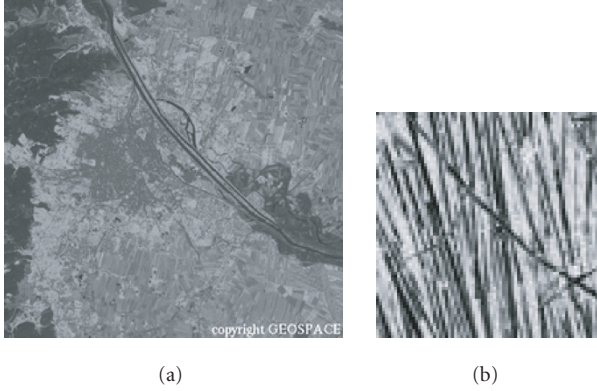


FIGURE 1: (a) Remote sensing image and (b) nature image.

Lakshmanan and Derin in [6]). First, a visit schedule $\{m_v\}$ as a function of v is established, where v denotes the time variable for this SA procedure. For each v , m_v identifies a component of the parameter vector θ . If $m_v = j$, then at time v , θ_j is updated as follows: a candidate value for θ_j is chosen at random between $\theta(v-1) - r$ and $\theta(v-1) + r$, for r appropriately small and where $\theta(v-1)$ denotes the value of θ_j before the update. This gives us a candidate parameter vector θ' . The following ratio is then computed with the candidate θ' and the old value $\theta(v-1)$:

$$\rho = \frac{[\prod_{s \in S} p(x_s | x_{N_s}, \theta')]^{1/T_0(v)}}{[\prod_{s \in S} p(x_s | x_{N_s}, \tilde{\theta}(v-1))]^{1/T_0(v)}} \quad (17)$$

$$= \sum_{s \in S} \exp \left(\frac{1}{T(v)} (\Delta E_1(\tilde{\theta}(v-1)) - \Delta E_2(\theta')) \right),$$

where $T_0(v)$ denotes the temperature in this SA procedure. Then $\tilde{\theta}(v)$ is chosen according to the following:

$$\tilde{\theta}(v) = \begin{cases} \theta' & \text{if } \rho > \varsigma, \\ \tilde{\theta}(v-1) & \text{otherwise,} \end{cases} \quad (18)$$

where ς is a random number with uniform distribution. This procedure generates a sequence $\{\tilde{\theta}(v)\}$ such that $\lim_{v \rightarrow \infty} \tilde{\theta}(v)$ maximizes pseudo-likelihood. ICM's implement is the same as SA, which may be regarded as SA with the extreme annealing schedule $T(n) = 0$.

The algorithm for the parameter estimation may now be stated explicitly as follows:

- (1) perform the image segmentation using initial parameters adopted by Pappas's adaptive MRF method and assuming $\gamma = 2$;
- (2) perform ICM to obtain coarse parameter estimation;
- (3) perform SA to obtain finer parameter estimation;
- (4) perform the image segmentation, and go to 3.

Simultaneous segmentation will be achieved as a by-product.

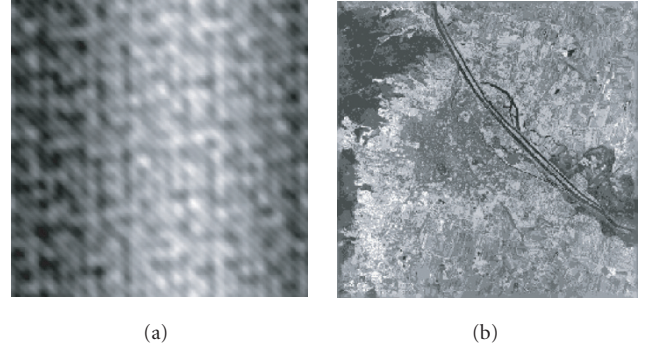


FIGURE 2: (a) Deterministic component and (b) pure random component.

6. EXPERIMENTAL RESULTS

The texture images used in this experiment is taken from Geospace. In the middle of the remote sensing image shown in Figure 2a, there is an urban area, while the other areas are suburban and mountainous areas. In many remote sensing applications, urban areas extraction is interesting. In the presented scale, urban and other areas all present texture characteristic, and so this is a complex scene segmentation problem. In the initial segmentation, selected parameters are defined as $\beta_s = 0, 5$; $K = 4$; $\sigma = 7$ gray levels; $\gamma = 2$, and the iteration number is 50. In fact, the final estimations are independent of initial values. Wold model decomposition is earlier than MRF segmentation and the Gaussian model is used to fit the data model. In the parameter estimation, the number of selected frequency points is 20, and the local maximum window is 5. To make simple, in our experiment, we use the homogenous MRF model including single and pairwise cliques. The edge of the image is processed in toroidal method. ICM-SA method is adopted. Temperature schedule is $T_2 = 1/(1/T_1 + 0.5)$, $T_1 = 100$, and the random value is $(\text{random}(1) - 0.5)$.

Figure 2 presents the results by using the Wold model decomposition: (a) presents a deterministic component, that is, a structural component, and it shows the texture periodicity and directionality and (b) presents a pure random component. We choose the deterministic component according to several main spectrum frequencies, as shown in Figure 3, which provide the predominant structure in the image (as discussed by Liu and Picard in [19]). Inverse transforming the component at these locations and scaling approximates the original image.

In Figures 4 and 5, the symbol $*$ denotes the last iteration results. The proportion between the single clique and the pairwise clique is denoted as γ , and β_1 and β_2 represent the diagonal potential and horizontal/vertical direction potential, respectively. It illustrates that the proportion is irrelevant to the potential parameters, and the changing beta is independent of the proportion γ . By SA, we can estimate the

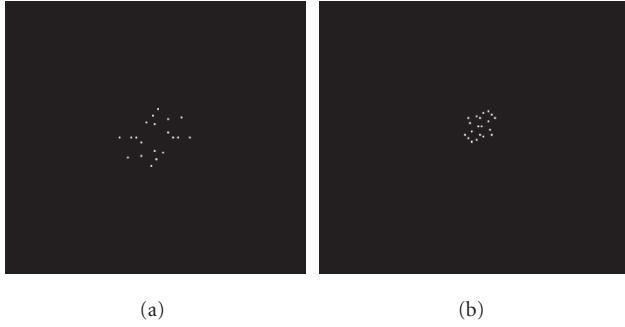


FIGURE 3: Main frequency spectrum of the deterministic part: (a) periodic spectrum; (b) directional spectrum.

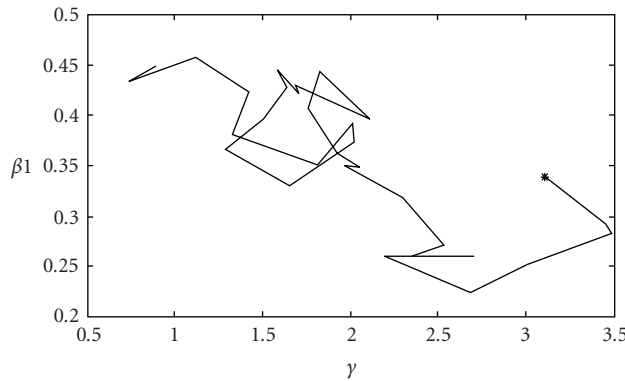


FIGURE 4: The relation between proportion γ and β_1 .

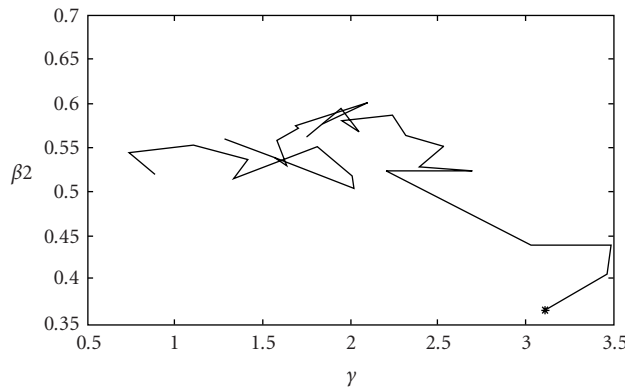


FIGURE 5: The relation between proportion γ and β_2 .

parameter. From the two figures, we can find that the best relative strength range of the two fields is $0.5 \sim 3.5$, so one can choose 2 as the initial value.

By choosing different values for γ , one can obtain different segmentation results, as in Figure 6. This is because the ratio of the two fields will influence the ability that the MRF model captures stochastic and structural texture components. Let a critical value be t , and if γ is bigger than t ,

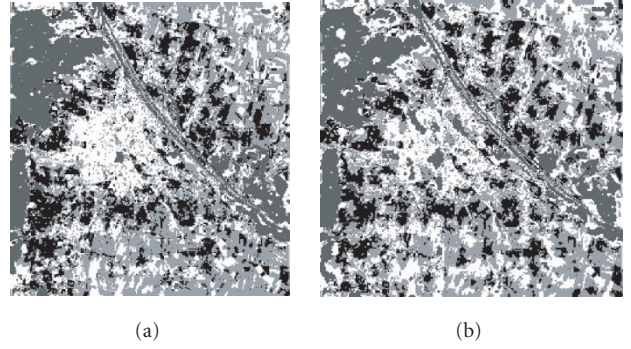


FIGURE 6: Segmentation results: (a) $\gamma = 1.8$ and (b) $\gamma = 2.8$.

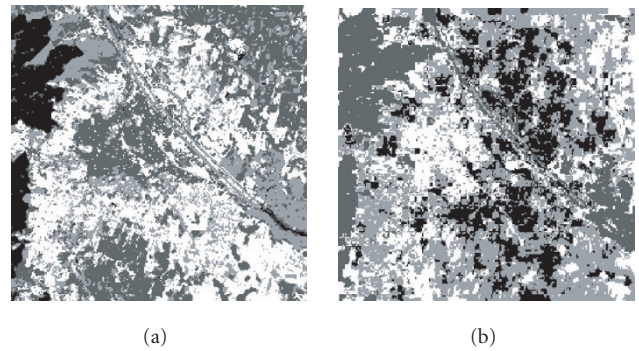


FIGURE 7: (a) MRF segmentation and (b) segmentation of the new algorithm.

then the segmented image will show obvious structural trait; by contrast, the segmented image have more stochastic trait. In Figure 7, (a) presents MRF-based pairwise segmentation only and (b) presents a result of the new algorithm. From Figure 7a, we can see that the segmentation has many errors, such as urban areas cannot be distinguished from the upper-right and bottom-left region, while there are fewer errors in (b) and it shows somewhat semantic characteristic.

Figure 8 illustrates the experiment results of an urban area against the other binary classifications. They correspond to Figures 7a and 7b, respectively. The pixels with white color represent urban areas while with dark color represent nonurban areas. Morphology postprocess may be needed in order to obtain better urban areas depiction.

We segment the image by using the new algorithm, given $K = 3$ and $K = 5$, respectively. In Figure 9a, the upper-left areas cannot be distinguished from the urban area. Figure 9b has the same good result as Figure 8b, but $K = 5$ takes more CPU time in our experiment.

In order to test the robustness of the new method, we consider another SPOT5 image. We wish to find the runway in Figure 10a, which is supposed to be the interesting object. Figure 10b is the segmentation using the new algorithm. One can observe that the runway is properly segmented.

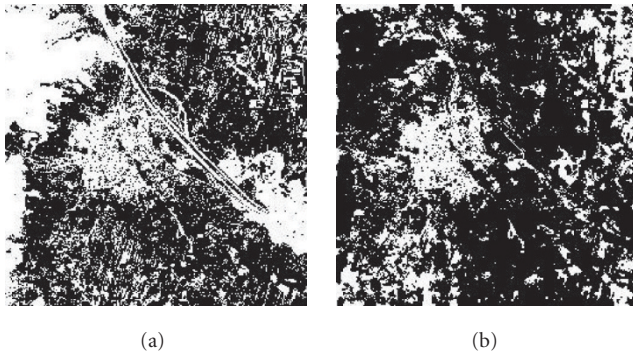
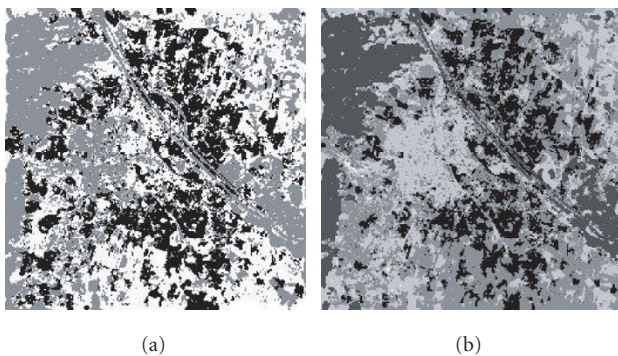
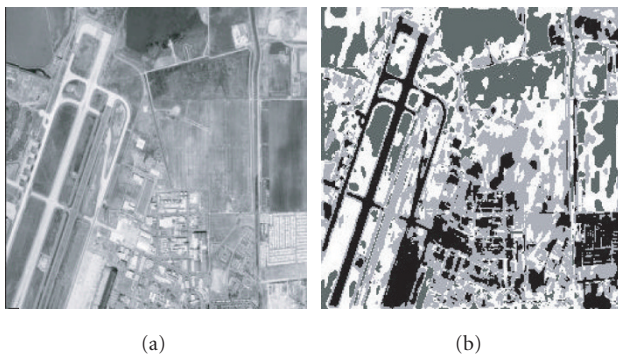


FIGURE 8: Urban area against the other binary classifications.

FIGURE 9: (a) Segmentation given $K = 3$ and (b) segmentation given $K = 5$.FIGURE 10: SPOT5 image segmentation given $K = 4$.

7. CONCLUSIONS

The usage of only pairwise in the MRF model can capture the stochastic component of texture, but not the structural. It is because the prior knowledge of the percentage of pixels in each region type is often unknown so that it is often assumed as 0 or equal, which produces a smoothed texture edge in the process of segmentation. This paper gives a new segmentation algorithm which simultaneously takes into account the stochastic and structural components of the texture by Wold

decomposition. As the decomposition can extract the texture structural component, we introduce it as the reference image of the external field in the MRF model. Due to the consideration of the texture structure, the resulting segmented image shows a semantic characteristic, which helps to understand the image better. In addition, a modified estimation procedure offers a simple and reliable scheme to model parameters.

ACKNOWLEDGMENT

The authors gratefully acknowledge Geospace for its image.

REFERENCES

- [1] C. Bouman and B. Liu, "Multiple resolution segmentation of textured images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 2, pp. 99–113, 1991.
- [2] D. K. Panjwani and G. Healey, "Markov random field models for unsupervised segmentation of textured color images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17, no. 10, pp. 939–954, 1995.
- [3] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 39–55, 1987.
- [4] M. L. Comer and E. J. Delp, "Segmentation of textured images using a multiresolution Gaussian autoregressive model," *IEEE Trans. Image Processing*, vol. 8, no. 3, pp. 408–420, 1999.
- [5] S. Yu, M. Berthod, and G. Giraudon, "Toward robust analysis of satellite images using map information-application to urban area detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 4, pp. 1925–1939, 1999.
- [6] S. Lakshmanan and H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 8, pp. 799–813, 1989.
- [7] T. N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Trans. Signal Processing*, vol. 40, no. 4, pp. 901–914, 1992.
- [8] X. Y. Yang and J. Liu, "Unsupervised texture segmentation with one-step mean shift and boundary Markov random fields," *Pattern Recognition Letters*, vol. 22, no. 10, pp. 1073–1081, 2001.
- [9] J. Zhang, W. Modestino, and D. A. Langan, "Maximum-likelihood parameter estimation for unsupervised stochastic model-based image segmentation," *IEEE Trans. Image Processing*, vol. 3, no. 4, pp. 404–420, 1994.
- [10] A. Lorette, X. Descombes, and J. Zerubia, "Texture analysis through a Markovian modelling and fuzzy classification: application to urban area extraction from satellite images," *International Journal of Computer Vision*, vol. 36, no. 3, pp. 221–236, 2000.
- [11] X. Descombes, M. Sigelle, and F. Preteux, "Estimating Gaussian Markov random field parameters in a nonstationary framework: application to remote sensing imaging," *IEEE Trans. Image Processing*, vol. 8, no. 4, pp. 490–503, 1999.
- [12] C.-S. Lu and P.-C. Chung, "Wold features for unsupervised texture segmentation," in *Proc. 14th IEEE International Conference on Pattern Recognition*, pp. 1689–1693, Brisbane, Australia, August 1998.
- [13] Y. Huang, K. L. Chan, and Z. H. Zhang, "Texture classification by multi-model feature integration using Bayesian networks," *Pattern Recognition Letters*, vol. 24, no. 1, pp. 393–401, 2003.

- [14] R. W. Picard, "Structured patterns from random fields," in *Proc. 26th IEEE Annual Asilomar Conference on Signals, Systems, and Computers*, vol. 2, pp. 1011–1015, Pacific Grove, Calif, USA, October 1992.
- [15] J. M. Francos, A. Z. Meiri, and B. Porat, "A unified texture model based on a 2-D Wold-like decomposition," *IEEE Trans. Signal Processing*, vol. 41, no. 8, pp. 2665–2678, 1993.
- [16] J. M. Francos, B. Porat, and A. Z. Meiri, "Orthogonal decompositions of 2-D nonhomogeneous discrete random fields," *Mathematics of Control, Signals and Systems*, vol. 8, no. 10, pp. 375–389, 1995.
- [17] Y. Wu and Y. Yoshida, "An efficient method for rotation and scaling invariant texture classification," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 4, pp. 2519–2522, Detroit, Mich, USA, May 1995.
- [18] R. Kashyap and R. Chellappa, "Estimation and choice of neighbors in spatial-interaction models of images," *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 60–72, 1983.
- [19] F. Liu and R. W. Picard, "Periodicity, directionality, and randomness: Wold features for image modeling and retrieval," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 7, pp. 722–733, 1996.

Feng Li was born in 1972. He received the B.E. degree in automatic control in 1996 from Nanchang University, China, and M.S. and Ph.D. degrees in automatic control in 1999 and 2003 from Gansu university of Technology and Huazhong university of Science and Technology, China, respectively. He is a Postdoctor at the Institute of Computing Technology, Chinese Academy of Sciences. His research interests are in the fields of image segmentation based on multirandom fields and artificial mobile terminal.



Jiaxiong Peng was born in 1934. He received the B.E. degree in automatic control in 1955 from Northeast University, China. He is a Professor at Huazhong University of Science and Technology. His research interests are in the fields of object recognition and image understanding.



Xiaojun Zheng was born in 1962. He received the B.E. and M.S. degrees in mechanics in 1983 and 1986 from Chinese National Defence University of Science and Technology, China, respectively, and Ph.D. degree in Intelligence Artificial in 1989 from Huazhong university of Science and Technology, China. He is a Professor at the Institute of Computing Technology, Chinese Academy of Sciences. His research interests are in the fields of wireless communication and artificial mobile terminal.

