# Adaptive Segmentation of Textured Images by Using the Coupled Markov Random Field Model

Yong Xia, *Student Member, IEEE*, (David) Dagan Feng, *Fellow, IEEE*, and Rongchun Zhao

***Abstract*—Although simple and efficient, traditional feature-based texture segmentation methods usually suffer from the intrinsical less inaccuracy, which is mainly caused by the oversimplified assumption that each textured subimage used to estimate a feature is homogeneous. To solve this problem, an adaptive segmentation algorithm based on the coupled Markov random field (CMRF) model is proposed in this paper. The CMRF model has two mutually dependent components: one models the observed image to estimate features, and the other models the labeling to achieve segmentation. When calculating the feature of each pixel, the homogeneity of the subimage is ensured by using only the pixels currently labeled as the same pattern. With the acquired features, the labeling is obtained through solving a maximum *a posteriori* problem. In our adaptive approach, the feature set and the labeling are mutually dependent on each other, and therefore are alternately optimized by using a simulated annealing scheme. With the gradual improvement of features' accuracy, the labeling is able to locate the exact boundary of each texture pattern adaptively. The proposed algorithm is compared with a simple MRF model based method in segmentation of Brodatz texture mosaics and real scene images. The satisfying experimental results demonstrate that the proposed approach can differentiate textured images more accurately.**

***Index Terms*—Image segmentation, image texture analysis, random field, simulated annealing.**

## I. Introduction

SEGMENTATION of textured images has long been an important and challenging topic in the image processing society. It aims to partition an image into several disjointed regions that are homogeneous with regards to some texture measures, so that subsequent higher level computer vision processing can be performed. According to how much a priori knowledge is involved, this problem can be divided into three subsets: supervised segmentation, semi-supervised segmentation, and unsupervised segmentation [1]. In this paper, we consider only the so-called semi-supervised problem, where the number of texture patterns is known but the information about their properties is not. During the past several decades, two types of algorithms have emerged as solution to this problem: model-based segmentation and feature-based segmentation.

In model-based algorithms, image segmentation is treated as an incomplete data problem, where the gray level of each pixel is known and the label, which designates the texture pattern the pixel belongs to, is missing. To solve this problem, the hierarchical two-level model [2], [3] is most commonly used. Each texture pattern in the observed image is modelled by a lower level spatial distribution to capture its property. The desired labels of all pixels are modelled by another higher level spatial distribution to characterize the blob-like region formation process [2]. By Means of sampling from the posterior distribution of the hierarchical model, simulated annealing may be used to estimate a suboptimal solution, which yields either the maximum *a posteriori* (MAP) [4]–[6] or the maximum posterior marginals (MPM) [7]. As a matter of fact, various forms of the Markov random field (MRF) [8] can be taken by the underlying components of the hierarchical model. Many simplifications can be adopted to make the posterior distribution computationally feasible [9], and a lot of model optimization and parameter estimation techniques can be applied to the model fitting [10]–[12]. Consequently, a variety of implementations have been presented throughout the literature [13]–[16]. However, none of those approaches have proved to be able to converge to a global minimum. No robust segmentation can be guaranteed unless the mechanism, which would allow an optimization process to escape local minimums, could be delicately established. Another disadvantage of those approaches is their complexity. Although all model parameters and the labelling can be estimated simultaneously, the alternate sampling process usually makes those approaches far more computationally intensive than their feature-based counterparts [1].

Feature-based segmentation algorithms can be briefly regarded as consisting of two successive processes: feature extraction and feature clustering. For each pixel, a feature vector is generated to indicate the gray-level statistics and local texture content over a window centered on that pixel. The range of features are diverse: model based statistics [17], [18], local spatial statistics [19], and statistics derived from the spatial-frequency domain [20]. Clustering algorithms which are applied to the acquired features can also take many different forms, such as statistical-based algorithms, neural network based algorithms, and various fuzzy algorithms [21]. However, traditional clustering methods may fail to fully utilize the spatial information provided by the feature set.

Manuscript received June 23, 2005; revised March 13, 2006. This work was supported in part by Grants from HK-RGC and ARC, in part by the National Science Foundation of China under Grant 60141002, and in part by the Aeronautical Science Foundation of China under Grant 02I53073. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zoltan Kato.

Y. Xia is with the School of Computers, Northwestern Polytechnical University, Xi'an 710072, China, and also with the School of Information Technologies, University of Sydney, NSW 2006, Australia (e-mail: yxia@it.usyd.edu.au; yxia@ieee.org).

D. Feng is with the School of Information Technologies, University of Sydney, NSW 2006, Australia, and also with the Center for Multimedia Signal Processing, Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong (e-mail: enfeng@de638.eie.polyu.edu.hk).

R. Zhao is with the School of Computers, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: rczhao@nwpu.edu.cn).

Digital Object Identifier 10.1109/TIP.2006.877513

In recent years, a lot of novel feature-based methods have been proposed, among which the algorithm developed by Deng and Clausi [22], [23] is shown to be successful. Since it is based on a simple MRF (SMRF) model, we call it the SMRF for convenience. The SMRF algorithm adopts the MAP-MRF framework [4], [24] and thus achieves segmentation by minimizing the weighted sum of two energy terms. One energy term is the sum of the Gaussian energy related to each feature, and another is the energy of the MRF characterizing the label distribution. With the first term alone, it turns out to be the separation of mixed data which results from several independent Gaussian distributions. It is the second term that imposes the smoothing and regularisation constraints on the segmentation result. Although highly simple and efficient, this approach, like all feature-based algorithms, has the drawback of being intrinsically less accurate, which is mainly attributed to the use of windows to calculate features. The estimated feature is not valid unless the texture in the estimation window is homogenous. It is usually assumed that the homogeneity is satisfied, whereas it is not true for boundary regions, where any estimation window may contain more than one texture patterns. The inhomogeneous texture will incur erroneous features, which may have a strong impact on the accuracy of the final segmentation. Obviously, the inaccuracy can only be avoided by using the pixels with the same label in each estimation window. Unfortunately, we won't know the exact label of each pixel until the optimal segmentation is obtained.

The primary contribution of this paper is to improve the accuracy of feature-based algorithms by proposing an adaptive segmentation of textured image. Our approach is based on the coupled Markov random field (CMRF) model, which has two mutually dependent components: a finite symmetric conditional Markov (FSCM) model [8] characterizes the observed image to obtain texture features [17], [18], and a multilevel Logistic (MLL) model [2] depicts the desired labelling to achieve MAP-MRF-based segmentation [4]–[6], [22], [23]. When computing the feature set, labelling is regarded as known so that, for each pixel, only the pixels with the same label inside an estimation window are used to calculate a feature. During segmentation, the obtained features are viewed as constant, and therefore the labelling can be estimated by solving the MAP problem, where features determine the likelihood probability and the MLL model imposes the smoothing and regularisation constraints. In our adaptive segmentation, the feature set and the labelling are alternately optimized by a simulated annealing scheme. By using the intermediate segmentation result, instead of an assumption, the homogeneity of the subimage which is used to estimate a feature is much improved. As a result, features gradually get more accurate and, accordingly, the boundary of each texture pattern in the labelling increasingly approaches the true one. Finally, this novel approach has been tested on mosaics of natural textures and the real scene images as well.

## II. Coupled Markov Random Field Model

The MRF theory can be traced back to the provocative work of Bosanov [25] in the 1960s. It was initially developed to model context dependent entities conveniently and consistently. Based on the significant contribution made by research in 1970s and 1980s [4], [25], [26], it has become a mathematically sound and computationally tractable tool in statistical image analysis. In this paper, an image is described by a hierarchical MRF model, which is detailed as follows.

An image is considered as a random field pair $\langle G, L\rangle$ defined on a $W \times H$ rectangular lattice $S = \{(i,j) : 1 \le i \le W, 1 \le j \le H\}$, which is indexed by the coordinate $(i,j)$. The gray-scale values are represented by $G = \{G_s = g_s : s \in S\}$, where $s = (s_i, s_j)$ denotes a specific site and the random variable $G_s$ takes integer values from the range [0, 255]. An observed image $g = \{g_s : s \in S\}$ is an instance of $G$. Similarly, the labels are denoted as $L = \{L_s = l_s : s \in S\}$, where the random variable $L_s$ takes value from an finite set $\Lambda = \{1, 2, \ldots, R\}$ and $R$ is the number of texture patterns appeared in the image. A labelling $l = \{l_s : s \in S\}$, which symbolizes a possible segmentation, is also called a configuration in the terminology of random field.

For the observed image $g$, let an estimation window with size $w \times w$ center on each site $s$. Within the window, all sites with the label $l_s$ form a set $W_s$, in which the subimage is assumed to be a MRF and described by a FSCM model $T^s$ with an associated second order neighborhood [8], shown as follows:

$$g_s = \sum_{t \in \eta_g} \theta_{s,t} \left(g_{s+t} + g_{s-t}\right) + e_s, s \in W_s^I \tag{1}$$

where $\eta_g = \{(0,1),(1,0),(1,1),(-1,1)\}$ is the set of shift vectors corresponding to the second order neighborhood system, $\{\theta_{s,t} : t \in \eta_g\}$ is the set of correlation coefficients associated with the set of translations from the central site $s$, $\{e_s\}$ is a stationary Gaussian noise sequence with variance $\sigma_s^2$, and $W_s^I$ is the interior subset of $W_s$.

The label field $L$ is also assumed to be a MRF and characterized by a second order multilevel logistic (MLL) model [2]. As suggested by the Hammersley-Clifford theorem [26], the joint distribution of all labels $p(L = l)$ follows a Gibbs distribution and takes the following form:

$$P(L = l) = Z^{-1} \exp\left\{-\frac{1}{T} E_L(l)\right\} \tag{2}$$

where $Z$ is a normalizing constant or partition function of the system, $T$ is a constant analogous to temperature, and $E_L(l)$ is the energy function.

Based on the FSCM model, a six-dimensional texture feature vector is defined as

$$f_s = \left(\mu_s, \hat{\sigma}_s^2, \hat{\theta}_{s,t} : t \in \eta_g\right) \tag{3}$$

where $\mu_s$ is the mean of the subimage, $\hat{\sigma}_s^2$ is the estimation of noise variance, and other four components $\hat{\theta}_{s,t}$ are the estimation of the correlation coefficients. An image can then be represented by a multidimensional random field $F = \{F_s = f_s;\ s \in S\}$, an instance of which is called a feature set $f = \{f_s : s \in S\}$. Traditionally, a feature vector is estimated by using the subimage within a regular window

without concerning the labels of those pixels. However, in our hierarchical model, only the pixels with a specific label in the estimation window are assumed to satisfy a FSCM model. That means the feature set $f$ is a function of both the gray-level image $g$ and the labelling , which can be expressed as

$$f = T(g, l). \tag{4}$$

As suggested by the MAP-MRF framework, the optimal labelling under current feature set $f$ can be obtained by maximizing the posterior probability of $L = l$ conditioned on $F = f$, indicated as follows:

$$l^*|_f = \max_{l \in \Omega_L} p(L = l | F = f) \tag{5}$$

where $\Omega_L$ is the configuration space, i.e., the set of all possible segmentations. According to the Bayes rule [28], the above equation can be further expressed as

$$l^*|_f = \max_{l \in \Omega_L} \frac{p(F = f | L = l)\, p(L = l)}{p(F = f)} \tag{6}$$

where $p(F = f)$ is the probability of feature set $f$, $p(L = l)$ is the prior probability of labeling $l$, and $p(F = f | L = l)$ is the likelihood probability of obtaining the feature set $f$ on condition of its neighborhood and label $l$.

From (4) and (6), it is obvious that the feature set and the labelling are mutually dependent in our CMRF model. As a result, segmentation of textured image can be formulated as the following optimization problem:

$$\begin{cases} f^* = T(g, l^*) \\ l^* = \max_{l \in \Omega_L} \frac{p(F = f^* | L = l) p(L = l)}{p(F = f^*)} \end{cases} \tag{7}$$

where $f^*$ and $l^*$ represent the optimal feature set and labelling, respectively.

## III. Adaptive Segmentation

In tradition feature-based segmentation methods, features can be computed first and segmentation is then performed. Unfortunately, the feature set and the labelling in our CMRF model, as shown in the previous section, mutually affect each other and cannot be obtained independently. Therefore, finding the globally optimal solution of the optimization problem given by (7) is computationally intractable. In this work, a step-wise optimization scheme is adopted to achieve an adaptive segmentation, in which the feature set and the labelling are alternatively updated. When calculating features, the labelling is assumed to be known and fixed. While refining the labelling, the features are assumed to be predetermined constants. Before describing the optimization technique, the calculation of features and the posterior probability is briefly described.

### A. Feature Estimation

Although parameters of the FSCM model $T^s$ can be estimated by different methods, the least-squares estimation proposed by Manjunath and Chellappa [17], [18] is adopted here because of its simplicity and consistency. Consequently, the feature vector of site $s$ is calculated as follows:

$$\hat{\Theta}_s = \left[ \sum_{s \in W_s^I} Q_s Q_s^T \right]^{-1} \left[ \sum_{s \in W_s^I} Q_s g_s \right] \tag{8}$$

$$\hat{\sigma}_s^2 = \frac{1}{|W_s^I|} \sum_{s \in W_s^I} \left[ g_s - \hat{\Theta}_s^T Q_s \right]^2 \tag{9}$$

$$\mu_s = \frac{1}{|W_s|} \sum_{s \in W_s} g_s \tag{10}$$

where $Q_s = \mathrm{col}[g_{s+t} + g_{s-t}; t \in \eta_g]$ is a column of gray-level and $\Theta_s = \mathrm{row}\left[\hat{\theta}_{s,t}; t \in \eta_g\right]$ is the parameter set. It should be mentioned that, when applying to the FSCM model, the gray level of each site should be subtracted by the local mean, since $E[g_s] = 0$, $s \in W_s$ is expected to meet the stationary assumption.

### B. Posterior Probability

With a known feature set $f$, the probability $p(F = l)$ does not vary with respect to any solution $L = l$ and can be eliminated. Consequently, (6) can be simplified as

$$l^*|_f = \max_{l \in \Omega_L} p(F = f | L = l)\, p(L = l). \tag{11}$$

Similar to the SMRF algorithm [22], [23], two assumptions are introduced to calculate above probabilities. One is that each random variable $F_s$ in the feature field $F$ obeys an independent Gaussian distribution with a different means vector $M_r$ and a covariance matrixe $\Sigma_r$, corresponding to the pattern $r$ site $s$ belongs to. The other is that only pair-wised cliques $\{s, t\}$ in the labelling field have non-zero potentials $V_{s,t}$. Furthermore, $V_{s,t} = -1$ if $l_s = l_t$, and $V_{s,t} = 1$ if $l_s \neq l_t$. The MAP problem given by (11) can then be solved by minimizing the following posterior energy:

$$l^*|_f = \min_{l \in \Omega_L} \left[ \beta \sum_{s \in S} E_G(f_s, l_s) + \sum_{s \in S} \sum_{t \in \eta_s} V_{s,t} \right] \tag{12}$$

where

$$E_G(f_s, l_s) = (f_s - M_{l_s})^T \Sigma_{l_s}^{-1} (f_s - M_{l_s}) + \ln\left[ (2\pi)^D |\Sigma_{l_s}| \right] \tag{13}$$

is the Gaussian energy, $D$ is the dimension of features, and $\beta$ is a weighting parameter balancing the contribution of the feature-related energy and the labelling-related energy to the

overall posterior energy. Slightly different from the SMRF algorithm, we don't assume that each component of the feature be independent on others. Consequently, the term $E_G(f_s, l_s)$ in (12) is the energy of a multivariable Gaussian function.

### C. Step-Wised Optimization

Generally, the energy function given by (12) is non-convex, and thus the minimization is limited by any deterministic algorithm. In our adaptive segmentation algorithm, the simulated annealing scheme [29], together with the Metropolis sampler [4], is adopted to achieve a step-wised optimization of the CMRF model.

Initially, a guess of labeling $l$ is given and a feature set $f$ is estimated under the assumption that the texture in each estimation window is homogeneous. In each iteration, labels are updated in a raster scan order. For each site $s$, a new labeling $l'$ is created by randomly choosing a new label for that site. That means those two configurations are almost the same except $l_s = r$ and ${l_s}' = r'$. Therefore, their energy can be evaluated by comparing only the following terms:

$$E_s = \beta E_G(f_s(r), r) + \sum_{t \in \eta_s} V_{s,t}|_{l_s = r} \tag{14}$$

$$E_s' = \beta E_G(f_s(r'), r') + \sum_{t \in \eta_s} V_{s,t}|_{l_s' = r'} \tag{15}$$

where $f_s(r)$ and $f_s(r')$ are the features corresponding to the assumptions of $l_s = r$ and ${l_s}' = r'$, respectively. If ${E_s}' < E_s$, the new labeling $l'$ will substitute for the old one, i.e., $l = l'$. Otherwise, the new labeling $l'$ will be accepted with a probability

$$p_s(n) = \exp\left(-\frac{{E_s}' - E_s}{T(n)}\right) \tag{16}$$

where $T(n)$ is the temperature coefficient used in simulated annealing scheme, and $n$ is the iteration number. Correspondingly, the feature set will also be updated according to the new labelling. The revised feature set is almost the same to the old one, except that the feature of site $s$ is given as follows:

$$f_s = \begin{cases} f_s(r), & l_s = r \\ f_s(r'), & l_s = r' \end{cases}. \tag{17}$$

The alternate updating of feature set and labelling will be performed iteratively until a stopping criterion is reached. For example, in our experiments, the optimization was terminated when the number of iteration arrives at 200, or when the labeling does not change any more.

During the optimization process, the mean vectors $M_r$ and the covariance matrix $\Sigma_r$ $(r \in \Lambda)$ of the feature set are used to calculate the Gaussian energy $E_G(f_s, l_s)$. Theoretically, each time when the labelling or feature set changes, those statistics should be re-estimated through the maximum likelihood estimation. However, for the sake of computational simplicity, the re-estimation has not been performed until the labels and features of all sites have been scanned once in our approach.

Another simplification is introduced to facilitate the calculation of the features $f_s(r)$ and $f_s(r')$. Taking the estimation of $f_s(r)$ for example, where the label of site $s$ is assumed to be $r$, only those pixels with label $r$ in the $w \times w$ estimation window are charactered by the FSCM model. The number of those pixels is denoted as $n_s(r)$. A large $n_s(r)$ means the texture in the window is primarily homogeneous, and a small $n_s(r)$ will definitely lead to a high estimation error. Therefore, only if $n_s(r)$ falls in the range $[n_{\min}, n_{\max}]$, the feature $f_s(r)$ will be re-estimated; otherwise, the original one will be used instead, i.e., $f_s(r) = f_s$.

Finally, it should be pointed out that the labelling is very critical for feature re-estimation, for only those pixels whose labels are the same will be described by a FSCM model and used to calculate the feature. However, in the early stage of the optimization, the labelling is far from optimal. Using such an inaccurate labelling to re-estimate features, the estimation error implied in the labelling will be transferred into the feature set and amplified through the iterative optimization steps. Therefore, feature es-estimation and updating will not be performed in the first $N_c$ iterations.

### D. Parameter Setting

The proposed segmentation algorithm involves many parameters. Although some parameters can be estimated automatically, others need to be derived empirically.

Parameter $w$ is the width of the square estimation window. To improve the accuracy of the lease square estimation used in the feature extraction, parameter $w$ should be as large as possible. Taking account of the computational expense, we choose $w$ as 21. Another parameter related to $w$ is the range $[n_{\min}, n_{\max}]$, which determines if re-estimation of a feature vector is needed. To determine this parameter, the impact of small subimage to the estimation accuracy should be investigated first. Three Brodatz textures (D9, D55 and D84) [30] and two natural textures (rough wall and sand) selected from the USC-SIPI Image Database [31] are used as five image samples, on each of which feature estimation has been performed 20 times by using different portion of pixels in a $20 \times 20$ estimation window. To equalize the contribution of different components, each dimension of those features has been normalized by its maximum value. The relative estimation error is defined as the Euclidean distance between the obtained feature and the feature acquired by using the entire window. For each texture, this experiment has been repeated ten times at randomly selected locations, and the average errors are depicted in Fig. 1. It is clear that the estimation is pretty accurate when more than 80% of pixels are used, while this portion drops below 40%, the error becomes relatively large. Taking both accuracy and efficiency into consideration, we empirically set $[n_{\min}, n_{\max}]$ be $[0.35w^2, 0.8w^2]$.

Parameter $N_c$ determines when the spatial information implied in the labelling begins to be used in feature optimization. Comparing with a small $N_c$, which may introduce inaccuracy to both the feature set and the labelling, a large $N_c$ appears to

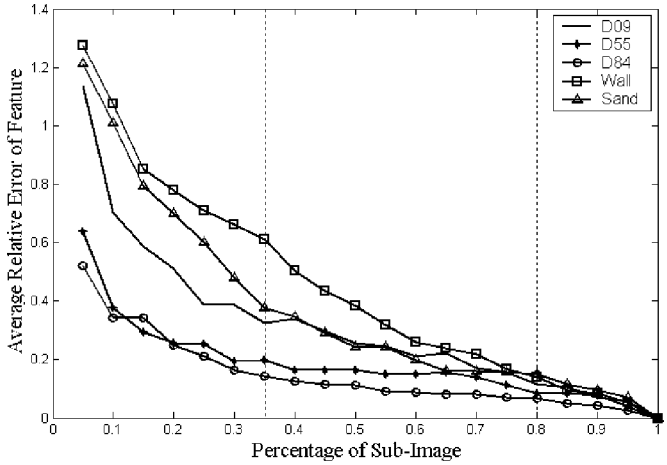


Fig. 1. Relationship between the average relative error of feature and the percentage of the subimage used in feature estimation.

be less harmful, though it will postpone the convergence of the system. In this study, we empirically set $N_c = 30$.

The energy weight $\beta$, which presents the tradeoff between the fidelity of the observed texture and the smoothness of the labelling, is very critical to the performance of this segmentation algorithm. It is expected to be relatively large in the early stage of the optimization so that the difference between features can play a dominant role, and considerably small in the late stage to introduce the spatial constraint. In our experiments, the following attenuation function used by the SMRF algorithm is adopted to acquire the variable weighting parameter:

$$\beta(n) = \beta_0 \cdot 0.95^n + \frac{1}{K} \tag{18}$$

where the initial weight $\beta_0$ is set as 13.

Finally, the annealing temperature sequence $T(n)$ plays a decisive role in the convergence of the simulated annealing. In each step, it determines the acceptable probability of high energy configuration. High temperature enables the optimization scheme to jump out of some local minimums, whereas low temperature ensures that the system will be frozen at certain low energy state. The optimal cooling scheme, which can assure a theoretical convergence, is computationally intractable [4]. Here, we use following temperature sequence as a suboptimal one:

$$T(n) = T_0 \cdot 0.95^n \tag{19}$$

where the initial temperature $T_0$ is set as 3.

## IV. Experimental Results

In this section, we present the results of our CMRF model based adaptive segmentation algorithm in three experiments. In order to assess its advantage in segmentation accuracy, the proposed approach is compared with the SMRF algorithm. With the purpose of making a fair comparison, two modifications have been made to the SMRF algorithm. First, the MRF model parameters mentioned in the previous section is adopted as features. Furthermore, the multivariable Gaussian distribution is used to compute the energy function, for the independence of each feature component can not be guaranteed. In our implementation of the SMRF algorithm, other parameters all take the suggested values given in [22]. As a result, the major difference between those two approaches is that the SMRF algorithm regards the feature set as known and unchangeable. Therefore, the following experiments mainly demonstrate the improvement of segmentation accuracy caused by adaptive process.

The first two comparative experiments have been performed on two sets of natural texture mosaics MII and MIV [32], which are generated by using 12 natural textures randomly chosen from the Brodatz album [30]. Those textures, together with their indexes and brief descriptions, are shown in Fig. 2. The first experiment tests the fundamental ability of our approach to distinguish two different texture patterns. Combining each of those 12 textures with every other, the set MII consists of ${}_{12}C_2 = 66$ samples, each of which is a mosaic of two textures with a size of $256 \times 256$ and a dynamic range of 256 gray levels. The Top row of Fig. 3 shows five example test cases (MII1 to MII5). The bottom row illustrates the results obtained by applying the proposed approach and the middle row gives the results of the SMRF algorithm. Table I compares the error percentage of the incorrectly classified pixels of the results given by Fig. 3. The edge of different regions in the original images is highlighted and those two texture patterns are indicated by black and white regions, respectively, in all results. It can be concluded from these results that, for such two class problem, our approach can achieve a successful segmentation with a higher accuracy, except for those sharp corners of texture regions, where the desired pattern cannot dominate the estimation window.

To demonstrate the advantage of the proposed algorithm in differentiating textures of more than two classes, the second experiment has been carried out on mosaics of four textures. Selecting four patterns from those 12 prototypes and combining them, the set MIV comprises ${}_{12}C_4 = 495$ samples. Five test cases, together with their corresponding segmentation results, are presented in Fig. 4. Similar to Fig. 3, the border of each texture region is drawn with white color in the original images and all segmentation results are shown by using arbitrarily selected gray level to highlight different regions. The segmentation errors of those five cases are given in Table II. It is apparent that our approach can achieve more accurate segmentation, especially in locating the edge of different patterns. Table III presents the average performance of the SMRF method and the proposed one on both databases. Although more mis-segmentation happened when dealing with four-class mosaics, the average error of our approach is less than half of that of the SMRF method. Meanwhile, it is shown in Table III that the standard deviation (Std) of segmentation errors of our method is smaller than that of the SMRF method, which means that our approach is more stable.

As mentioned before, the labeling optimization is dominated by the texture features in the early stage. When the energy weight $\beta$ is large, the effect of the prior energy can almost be ignored and the problem is somehow equal to the classification of mixed Gaussian patterns by using maximum likelihood estimation. In this case, the initial guess of labeling may have a strong impact on the final convergence. If the modes of all patterns are well separated, as in the first experiment, a

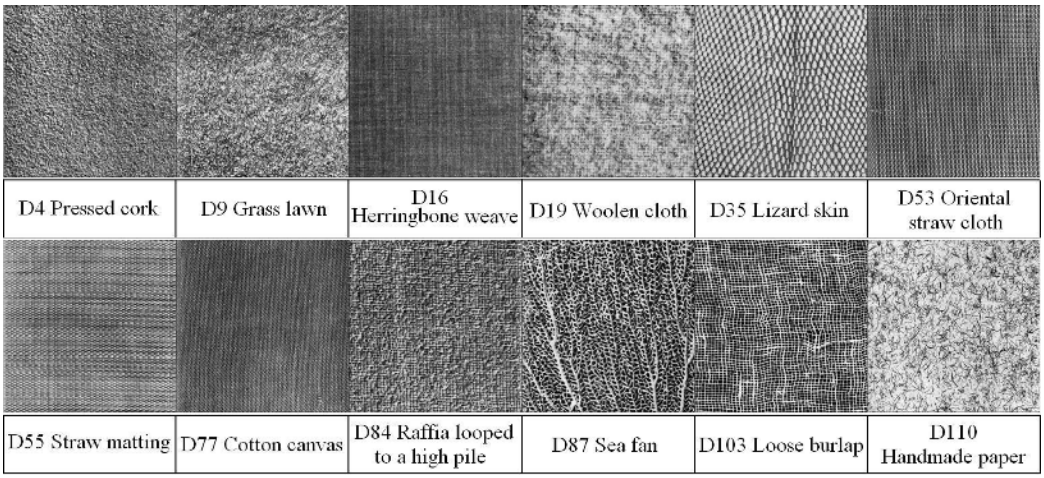


Fig. 2. Twelve natural textures from Brodatz album. The index and a brief description are shown under each texture.

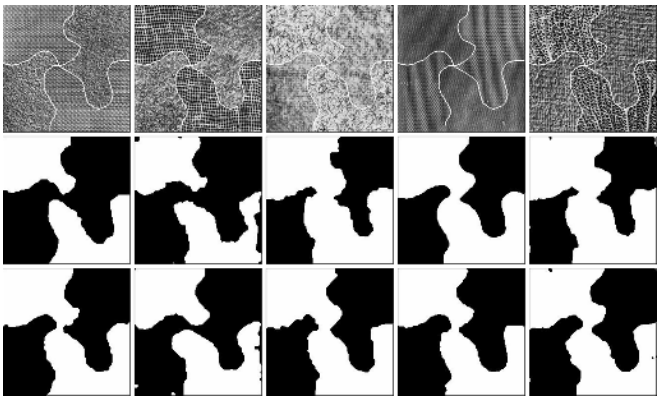

Fig. 3. Five test cases of mosaics of two textures (MII1 – MII5) and their segmentations by applying (middle row) the SMRF method and (bottom row) the proposed method.

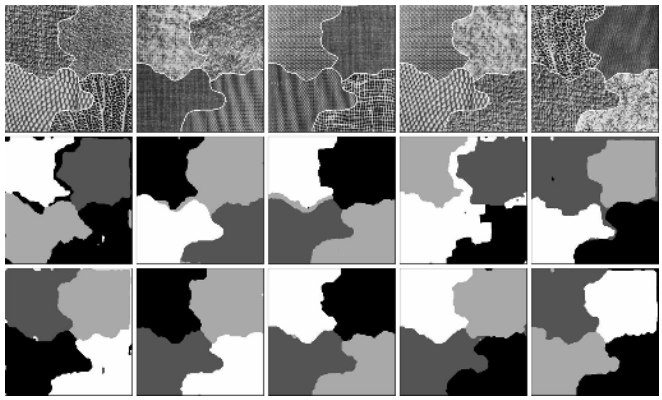

Fig. 4. Five test cases of mosaics of four textures (MIV1–MIV5) and their segmentations by applying (middle row) the SMRF method and (bottom row) the proposed method.

TABLE I
ERROR PERCENTAGE OF MISCLASSIFIED PIXELS ON IMAGE MII1–MII5

| Image Index | Texture Components | Methods | |
|---|---|---|---|
| | | SMRF | Proposed |
| MII1 | D4 vs. D55 | 6.37 % | 1.37 % |
| MII2 | D9 vs. D103 | 7.32 % | 3.49 % |
| MII3 | D19 vs. D110 | 8.17 % | 1.86 % |
| MII4 | D53 vs. D77 | 2.62 % | 1.80 % |
| MII5 | D84 vs. D87 | 9.57 % | 2.25 % |

TABLE II
ERROR PERCENTAGE OF MISCLASSIFIED PIXELS ON IMAGE MIV1–MIV5

| Image Index | Texture Components | Methods | |
|---|---|---|---|
| | | SMRF | Proposed |
| MIV1 | D4 vs. D35 vs. D84 vs. D87 | 9.70 % | 3.76 % |
| MIV2 | D9 vs. D16 vs. D19 vs. D53 | 4.97 % | 2.27 % |
| MIV3 | D16 vs. D53 vs. D55 vs. D103 | 5.65 % | 1.56 % |
| MIV4 | D19 vs. D35 vs. D55 vs. D84 | 10.73 % | 2.10 % |
| MIV5 | D77 vs. D84 vs. D87 vs. D110 | 8.10 % | 3.00 % |

TABLE III
AVERAGE ERROR PERCENTAGE OF MISCLASSIFIED PIXELS

| Image Dataset | Size of Database | Error of SMRF | | Error of Proposed | |
|---|---|---|---|---|---|
| | | Mean | Std | Mean | Std |
| MII | 66 images | 5.69 % | 0.0283 | 2.38 % | 0.0101 |
| MIV | 495 images | 12.32 % | 0.0907 | 5.84 % | 0.0743 |

randomly selected initial labeling may lead to an acceptable result. Otherwise, a better initialization is needed. With the increasing of patterns in the image, it becomes tremendously difficult to distinguish the features of each texture in the second experiment. To obtain a reasonable initialization, one sixteenth of the original features are clustered by the well-known Fuzzy C-Mean (FCM) algorithm [33]. Half labels are initialized according to the clustering results to increase the probability of correct convergence, and the other labels are randomly initialized to avoid a local minimum. Regardless of this, there are still 57 failed test cases (11.5%) in set MIV, which explains why the average segmentation error over the entire set is much higher than that of the given samples.

Next, we apply the presented comparative experiment to a number of natural images. The segmentations for a picture of sea and beach, a grassland scene, a perspective of buildings, and a portrait of a lizard are shown in Fig. 5. The original images are given in the first row, and the results obtained by the SMRF method and the proposed algorithm are shown in the

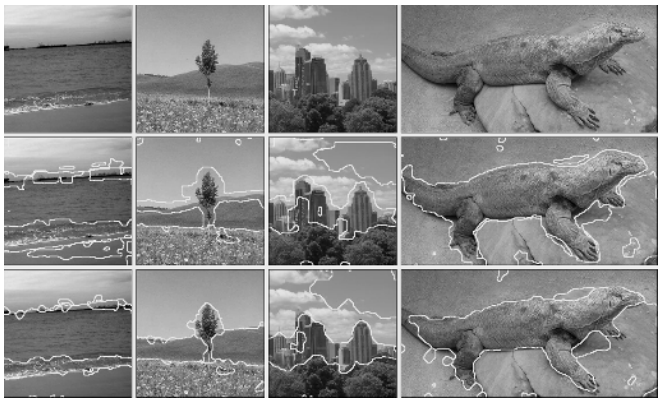

Fig. 5. Four real scene images and their segmentations by applying (middle row) the SMRF method and (bottom row) the proposed method.

TABLE IV
AVERAGE TIME COST OF TWO SEGMENTATION METHODS

| Images Dataset | SMRF | Proposed |
|---|---|---|
| MII (66 samples) | 38.766 s | 53.061 s |
| MIV (495 samples) | 43.860 s | 52.563 s |

middle and bottom row, respectively. The segmentation of textures in those images is evaluated subjectively. For a better comparison of those results, the border of each segmented region is directly drawn and overlaid on the original images. Compared with the results of the SMRF method, the segmented regions of our approach appear to agree better with the regions we would perceived as distinct, if we try not to make use of semantics like "sky," "beach," "clouds," "buildings," "grass," "trees," etc. Although some under segmentation occurs when applying our algorithm to the test samples shown in the right two columns of Fig. 5, over segmentation appears more severe in the corresponding results of the SMRF method. The proposed segmentation algorithm, on the whole, can detect boundaries of different textured regions more exactly in all four test cases. This conclusion is completely in accordance to the results reported in the previous experiments.

Although more accurate, the proposed approach has higher computational complexity than conventional approaches because of the re-estimation of some features in the segmentation process. During the optimization, the edge of different regions moves gradually and slowly. Compared with the size of estimation window, the changing edge in each step is unapparent. Therefore, it is not necessary to calculate the feature of a given site in every iteration. In our implementation, once the feature of a site is updated, it will not be recalculated in the next $\lfloor w/4 \rfloor$ iterations. This means a lot of computation has been saved. Table IV gives the average time cost of both methods in the first two experiments (Intel Pentium III 871-MHz processor and 512-M memory). It is revealed that the complexity of our approach is only slightly higher than that of the SMRF method.

## V. CONCLUSION

In this work, an adaptive algorithm based on the coupled MRF model is proposed to achieve accurate segmentation of textured images. By using the intermediate segmentation result, the accuracy of feature extraction is much improved. With the refined feature set, the segmentation result also gets better. Comparative experiments have proved the success of the proposed approach.

In the proposed CMRF model, the features of each pattern are assumed to satisfy a Gaussian distribution. This assumption usually simplifies the expression of likelihood probability. However, it may be a major cause of this algorithm's failure on some images, where the Gaussian distribution may not be an appropriate choice. This problem will be further investigated in future. Furthermore, although this work focuses on MRF-based features, the proposed idea of simultaneously refining features and labels may provide some inspiration to other feature-based texture segmentation approaches. In fact, the feature used in this paper has its limitation to model textures with strong edges/directionality or large-scale primitives. Additional future work is to explore different texture features to segment various images.

## ACKNOWLEDGMENT

The authors would like to thank Dr. H. W. Deng, Associate Prof. D. Clausi, University of Waterloo, and Prof. Y. N. Zhang, Northwestern Polytechnical University, for their invaluable help and support. The authors would also like to thank the Associate Editor and the anonymous reviewers for their insightful and constructive comments, which have greatly helped to improve both the contents and the representation of this paper.

## REFERENCES

[1] S. A. Barker, "Image Segmentation Using Markov Random Field Models," Dept. Eng., Cambridge Univ., Cambridge, U.K., 1998, Ph.D. dissertation.
[2] S. Z. Li, , C. R. Rao and D. N. Shanbhag, Eds., "Modeling image analysis problems using Markov random fields," in *Stochastic Processes: Modeling and Simulation*. New Providence, NJ: Elsevier Science, 2000, vol. 20, Handbook of Statistics, pp. 1–43.
[3] Z. Kato, M. Berthod, and J. Zerubia, "A hierarchical Markov random field model and multi-temperature annealing for parallel image classification," *CVGIP: Graph. Models and Image Process.*, vol. 58, pp. 18–37, Jan. 1996.
[4] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. MPAI-6, no. 6, pp. 721–741, Nov. 1984.
[5] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. MPAI-9, no. 1, pp. 39–55, Jan. 1987.
[6] C. S. Won and H. Derin, "Unsupervised segmentation of noisy and texture images using Markov random field," *Comput. Vis. Graph. Image Process.*, vol. 54, no. 4, pp. 308–328, Jul. 1992.
[7] M. L. Comer and E. J. Delp, "Segmentation of textured images using a multiresolution Gaussian autoregressive model," *IEEE Trans. Image Process.*, vol. 8, no. 3, pp. 408–420, Mar. 1999.
[8] R. Kashyap and R. Chellappa, "Estimation and choice of neighbors in spatial interaction models of images," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 1, pp. 60–72, Jan. 1983.
[9] D. E. Melas and S. P. Wilson, "Double Markov random fields and Bayesian image segmentation," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 357–365, Feb. 2002.
[10] J. Besag, "On the statistical analysis of dirty pictures," *J. Royal Statist. Soc. B*, vol. 48, no. 3, pp. 259–302, 1986.
[11] P. Andrey and P. Tarroux, "Unsupervised segmentation of Markov random field modeled textured images using selectionist relaxation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 252–262, Mar. 1998.
[12] S. A. Barker, "Unsupervised image segmentation using Markov random field models," *Pattern Recognit.*, vol. 33, pp. 587–602, 2000.
[13] F. C. Jeng and J. W. Woods, "Compound Gauss-Markov random fields for image segmentation," *IEEE Trans. Signal Process.*, vol. 39, no. 3, pp. 683–697, Mar. 1991.

[14] S. Krishnamachari and R. Chellappa, "Multiresolution Gauss-Markov random field models for texture segmentation," *IEEE Trans. Image Process.*, vol. 6, no. 2, pp. 251–267, Feb. 1997.
[15] A. Sarkar, M. K. Biswas, and K. M. S. Sharma, "A simple unsupervised MRF model based image segmentation approach," *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 801–812, May 2000.
[16] F. Yang and T. Jiang, "Pixon-based image segmentation with Markov random fields," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1552–1559, Dec. 2003.
[17] B. S. Manjunath, T. Simchony, and R. Chellappa, "Stochastic and deterministic networks for texture segmentation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 6, pp. 1039–1049, Jun. 1990.
[18] B. S. Manjunath and R. Chellappa, "Unsupervised texture segmentation using Markov random fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 5, pp. 478–482, May 1991.
[19] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, no. 5, pp. 796–804, May 1979.
[20] D. Dunn and W. E. Higgins, "Optimal gabor filters for texture segmentation," *IEEE Trans. Image Process.*, vol. 4, no. 7, pp. 947–964, Jul. 1995.
[21] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, pp. 264–323, 1999.
[22] H. Deng and D. A. Clausi, "Unsupervised image segmentation using a simple MRF model with a new implementation scheme," *Pattern Recognit.*, vol. 37, no. 12, pp. 2323–2335, 2004.
[23] ——, "Unsupervised segmentation of synthetic aperture radar sea ice imagery using a novel Markov random field model," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 528–538, Mar. 2005.
[24] R. C. Dubes and A. K. Jain, "Random field models in image analysis," *J. Appl. Statist.*, vol. 16, no. 2, pp. 131–164, 1989.
[25] Y. A. Rosanov, "On Gaussian fields with given conditional distributions," *Theory Probab. Appl.*, vol. XII, no. 3, pp. 381–391, 1967.
[26] J. M. Hammersley and P. Clifford, *Markov Field on Finite Graphs and Lattices*, 1971, unpublished.
[27] J. Besag, "Spatial interaction and the statistical analysis of lattice systems (with discussion)," *J. Roy. Statist. Soc. B*, vol. 36, pp. 192–236, 1974.
[28] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001.
[29] S. Kirkpatrick, C. D. Gelatt, Jr, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
[30] P. Brodatz, *Texture: A Photographic Album for Artists and Designers*. New York: Dover, 1966.
[31] The USC-SIPI Image Database Viterbi School of Engineering, Univ. of Southern California, Los Angeles [Online]. Available: http://sipi.usc.edu/database
[32] Y. Xia, (D.) D. Feng, and R. Zhao, "Morphological multifractal estimation for image segmentation," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 614–623, Mar. 2006.
[33] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.

**Yong Xia** (S'05) received the B.E. and M.E. degrees in computer applications in 2001 and 2004, respectively, from Northwestern Polytechnical University, Xi'an, China, where he is currently pursuing the Ph.D. degree in the School of Computers.

During 2003, he was a Research Assistant in the Center for Multimedia Signal Processing, Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong. From 2004 to 2006, he was a Visiting Researcher in the School of Information Technologies, University of Sydney, Sydney, NSW, Australia. His research interests include image analysis, multimedia signal processing, pattern recognition, and machine learning.

Mr. Xia is a student member of the Chinese Society of Image and Graphics (CSIG).

**(David) Dagan Feng** (S'88–M'88– SM'94–F'03) received the Ph.D. degree in computer science from the University of California at Los Angeles (UCLA) in 1988.

After briefly working as an Assistant Professor in U.S., he joined the University of Sydney, Sydney, NSW, Australia, as a Lecturer, Senior Lecturer, Reader, and then Professor. He is the former Head of Department of Computer Science/School of Information Technologies. He is currently the Associate Dean of the Faculty of Science and Director of the Biomedical and Multmedia Information (BMIT) Research Group at the University of Sydney, and Professor and Deputy Director, Center for Multimedia Signal Processing, Department of Electronic and Information Engineering, Hong Kong Polytechnic University. He has published over 300 scholarly research papers, pioneered several new research directions, and has made a number of landmark contributions in his field. Many of his research results have been translated into solutions to real-life problems worldwide and have made tremendous improvements to the quality of life.

Dr. Feng is currently Special Area Editor of the IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, Advisor for the *International Journal of Image and Graphics*, Chairman of IFAC-TC-BIOMED, and Chairman of the International Program and National Organizing Committees for the IFAC 2003 Symposium. He received the Crump Prize for Excellence in Medical Engineering. He is a Fellow of ACS, HKIE, and IEE.

**Rongchun Zhao** received the M.E. degree in weapon control and air force engineering from the PLA's Institute of Military Engineering, Harbin, China, in 1960.

He was a Senior Visiting Scholar in the Department of Electric and Electronics, Surrey University, Surrey, Guildford, U.K., from 1989 to 1990. Currently, he is a Professor with the School of Computer, Northwestern Polytechnical University, Xi'an, China. He has been Head of the School of Computers, and is currently a member of the Academic Committee of Northwestern Polytechnical University. His research interests include speech processing, image analysis and comprehension, computer vision, virtual reality and pattern recognition. He has published more than 100 scholarly research papers and two monographs. He is Founder and Director of the Provincial Key Laboratory of Speech and Image Information Processing (SIIP), Northwestern Polytechnical University. He has also been appointed as Vice President of the China Society of Image and Graphics, Vice President of the China Society of Stereology, Vice Director of the Signal Processing section of the Chinese Institute of Electronics, and President of Shaanxi Signal Processing Association.