Classification of High Spatial Resolution Imagery Using Improved Gaussian Markov Random-Field-Based Texture Features

Yindi Zhao, Liangpei Zhang, Member, IEEE, Pingxiang Li, Member, IEEE, and Bo Huang

Abstract—Gaussian Markov random fields (GMRFs) are used to analyze textures. GMRFs measure the interdependence of neighboring pixels within a texture to produce features. In this paper, neighboring pixels are taken into account in a priority sequence according to their distance from the center pixel, and a step-by-step least squares method is proposed to extract a novel set of GMRF texture features, named as PS-GMRF. A complete procedure is first designed to classify texture samples of QuickBird imagery. After texture feature extraction, a subset of PS-GMRF features is obtained by the sequential floating forward-selection method. Then, the maximum a posteriori iterated conditional mode classification algorithm is used, involving the selected PS-GMRF texture features in combination with spectral features. The experimental results show that the performance of classifying texture samples on high spatial resolution QuickBird satellite imagery is improved when texture features and spectral features are used jointly, and PS-GMRF features have a higher discrimination power compared to the classical GMRF features, making a notable improvement in classification accuracy from 71.84% to 94.01%. On the other hand, it is found that one of the PS-GMRF texture features—the lowest order variance—is effective for residential-area detection. Some results for IKONOS and SPOT-5 images show that the integration of the lowest order variance with spectral features improves the classification accuracy compared to classification with purely spectral features.

Index Terms—Classifying texture samples, Gaussian Markov random fields (GMRFs), least squares (LS) method, priority sequence, residential-area detection.

I. INTRODUCTION

SINCE the successful launch of high-resolution sensors and the wide application of high spatial resolution satellite data, it has been possible to monitor environmental changes on a small spatial scale. However, improved spatial resolution does not always lead to better classification results when employing conventional spectral classification methods [1]. While increases in spatial resolution create increased amounts of information detail, this also creates higher spectral variance within each class corresponding to land-cover units, which decreases their spectral separability and results in lower classification accuracy. Texture is a very useful property of spatial-structure information in high-resolution images [2]. By incorporating texture information in the classification, heterogeneous objects and different objects with the same spectral characteristics can be detected effectively, and higher classification accuracies can be obtained.

Texture analysis, which provides a complementary tool to the interpretation of high-resolution satellite images, has received great attention in image processing. Over the last two decades, many approaches to texture feature extraction have been developed, categorized roughly as feature-based methods such as the gray-level cooccurrence matrix, model-based methods such as Markov random-field (MRF) models, and structural methods [3]. Comparison and fusion of different types of texture features can be found in the literature [4], [5].

This paper focuses on extracting texture features based on MRF models. These models characterize spatial interactions between the central pixel of a texture image and its neighboring pixels to describe the local structure of an image [6]. Early work on texture classification using MRF was done by Chellappa and Chatterjee [7], wherein texture features were extracted based on the Gaussian MRF (GMRF), and a simple minimum-distance classifier was used for the texture image classification. More systematic theories about satellite image analysis were built up using MRF [8], [9] with emphasis on the application of MRF for spatial information extraction from satellite images. Torres-Torriti and Jouan [10] evaluated the performance of MRF models and Gabor filters for computing texture properties, which were then used as discrimination features to classify different land-cover types in synthetic aperture radar (SAR) imagery. Clausi and Deng [11] demonstrated that both texture and tonal features are necessary to properly perform segmentation and classification of SAR sea-ice imagery. Subsequently, Clausi and Yue compared MRF and gray-level co-occurrence probability (GLCP) for texture analysis of SAR sea-ice imagery [12].

However, the most relevant research about MRF regards all the neighboring pixels equally and puts the whole neighbor pixels together into the corresponding MRF model equation [7]–[12]. In this paper, the neighboring pixels will be treated with a priority sequence, in accordance with the distance from the center pixel, and a new set of MRF texture features will be computed based on this consideration, followed by feature...
selection to reduce data dimensions and increase computational efficiency. The objectives of this paper are twofold: to assess the potential of new MRF texture features applied to classification of high spatial resolution imagery and to assess the effect of integrating spectral features with texture features in high-resolution-imagery classification.

The rest of this paper is organized as follows. Section II describes the novel MRF texture features. Section III briefly illustrates the classification method. Section IV provides some comparison experiments using QuickBird samples. Detecting residential areas using the new MRF features is studied in Section V, and conclusions are given in Section VI.

II. TEXTURE FEATURE EXTRACTION

A texture feature extraction method is a computational method that analyzes the pixels contained in a certain window of an input image and generates a set of values that represents the contents of that window. An ideal texture feature extraction method should be capable of generating different values for different textures. One important property of an MRF model is that it uses a finite number of parameters to characterize spatial interactions of pixels to describe an image region. A typical MRF model is the GMRF model, which is widely used to model image textures by utilizing spectral as well as spatial information.

A. Novel Texture Features Based on GMRF Models

In order to make this paper self-sufficient, some important concepts for GMRF models are restated here. Markovianity refers to the fact that the central pixel of an image interacts with only the neighboring pixels, independent of other pixels [6]. A GMRF model specifies spatial dependencies between the central pixel and its neighboring pixels by defining a local conditional probability distribution, which is assumed to be Gaussian [13]. The number and position of neighboring pixels are determined by the neighbor order defined using a Euclidean distance [14]. Up to fifth-order neighbors are depicted in Fig. 1; wherein the number indicates the neighbor order to which the underlying site belongs. Note that one neighbor set with order 1 includes all the neighbors of sets with order one to 1. For example, a fourth-order GMRF model centered on an image cell s, would include those cells marked one to four, and its corresponding neighbor set can be denoted as a set of shift vectors $\mathbf{RN} = \{(0, 1), (1, 0), (1, -1), (1, 1), (0, 2), (2, 0), (1, -2), (1, 2), (2, -1), (2, 1)\} \cup \{(0, -1), (-1, 0), (-1, 1), (-1, -1), (0, -2), (-2, 0), (-1, 2), (-1, -2), (-2, 1), (-2, -1)\}$.

Let $S = \{s = (i, j)|0 \leq i \leq M - 1, 0 \leq j \leq N - 1\}$ denote the set of grid points in the $M \times N$ lattice corresponding to the pixels in the image region. Supposing that the image is modeled as a fourth-order GMRF model $f(s)$, the gray-level intensity of a pixel $s$ has a local conditional probability density function

$$p(f(s)|f_{\mathbf{RN}}(s)) = \frac{1}{\sqrt{2\pi\nu}} \exp \left\{-\frac{1}{2\nu} \left( f(s) - \mu - \sum_{r \in \mathbf{RN}} \beta(r)(f(s + r) - \mu) \right)^2 \right\}$$

where $f_{\mathbf{RN}}(s) = \{f(s + r)|r \in \mathbf{RN}\}$ stands for the set of values at the neighboring site of the central pixel $s$, $\mu$ is the mean gray value of the whole image, $\beta(r)$ are the model parameters, and $\nu$ is the conditional variance. On the other hand, the gray value of the central pixel $s$, $f(s)$ is also represented as a linear combination of values of its neighbor pixels and an additive noise [6], obeying the following difference equation:

$$f(s) - \mu = \sum_{r \in \mathbf{RN}} \beta(r) (f(s + r) - \mu) + e(s)$$

where $\{e(s)\}$ is a zero-mean Gaussian noise sequence with the correlation structure

$$\text{Cov}[e(s), e(s + r)] = \begin{cases} \nu, & r = (0, 0) \\ -\beta(r)\nu, & r \in \mathbf{RN} \\ 0, & \text{otherwise}. \end{cases}$$

Since the power spectrum associated with (2) must be real and positive [15], there are $r \in \mathbf{RN} \Leftrightarrow -r \in \mathbf{RN}$ and $\beta(r) = \beta(-r)$. Moreover, (2) can be rewritten by

$$f(s) - \mu = e(s) + \sum_{r \in \mathbf{RN}} \beta(r)[(f(s + r) - \mu) + (f(s - r) - \mu)]$$

where $\mathbf{RN}$ is an asymmetric neighbor set such that if $r \in \mathbf{RN}$, then $-r \notin \mathbf{RN}$ and $\mathbf{RN} = \{r|r \in \mathbf{RN}\} \cup \{|r| - r \in \mathbf{RN}\}$.

The parameters $\beta(r)$ and conditional variance $\nu$ describe the GMRF models and characterize textures. When they are unknown, they should be calculated. There are many existing methods for estimating those unknowns but none can guarantee consistency\(^1\) as well as stability.\(^2\) The covariance matrix in the expression for the joint probability density of MRF must be positive definite.

\(^1\)The estimates converge toward the true values of the parameters.

\(^2\)The covariance matrix in the expression for the joint probability density of MRF must be positive definite.
where \( Q(s) = \text{col}[(f(s + r) - \mu) + (f(s - r) - \mu) | r \in \mathbb{R}^N] \). The estimate \( \hat{\nu} \) of the conditional variance \( \nu \) is calculated by

\[
\hat{\nu} = \frac{1}{M \times N} \sum_{s \in S} [(f(s) - \mu) - Q(s)T \hat{\beta}]^2.
\]

A general GMRF-based approach employs \( \eta = \{ \hat{\beta}, \hat{\nu} \} \) as a texture feature vector. The classical GMRF (CGMRF)-based method thinks that all of the neighboring pixels, which are treated equally, interact on the center pixel simultaneously. However, it is reasonable that neighboring pixels have influence on the center pixel in a priority sequence, i.e., the closer the neighboring pixel to the center pixel, the higher the priority. To study this situation, a step-by-step LS (SSLS) method is proposed, to derive a different set of GMRF texture features. For this purpose, the parameters of the GMRF model can be divided into a certain number of groups and then estimated step-by-step, adopting the following rules.

Rule 1) The number of groups is equal to the order of the GMRF model.

Rule 2) The parameters for neighbors with different distances from the center pixel should be divided into different groups.

Rule 3) The parameters for neighbors closer to the center pixel should be estimated before those for neighbors further from the center pixel.

Note that the purpose is not to generate facsimiles of the texture but to assess a discrimination measure among different textures. For a fourth-order GMRF model, \( \beta \) is divided into four groups: \( \beta^{(1)}, \beta^{(2)}, \beta^{(3)}, \) and \( \beta^{(4)} \). Each group \( \beta^{(k)} \) corresponds to a group of neighbors, denoted by \( \mathbb{R}^N(k) \), that is, \( \mathbb{R}^N(1) = \{(0,1), (1,0)\}, \mathbb{R}^N(2) = \{(1,-1), (1,1)\}, \mathbb{R}^N(3) = \{(0,2), (2,0)\}, \) and \( \mathbb{R}^N(4) = \{(1,-2), (2,1)\} \). For the sake of simplicity, let us denote the original GMRF features as “classical GMRF” features, and denote the new GMRF features that consider the priority sequence, computed using the SSLS estimate method, as PS-GMRF features. A PS-GMRF feature vector was constructed using the results of each estimating step, namely, \( \psi = \{\beta^{(1)}, \beta^{(2)}, \ldots, \beta^{(m)}\} \). By means of SSLS, the lower order parameters are independent of those of higher orders. As a result, it is not necessary to determine the exact order of the GMRF model for a certain texture, and a higher order model is preferred. For some texture, if the order of the model used is higher than the exact order, the correlations between the center pixel and the higher order neighbor pixels are very weak, and the higher order parameters, which would not play an important role in texture recognition, will be abandoned by means of feature selection (described in Section II-B). The proposed SSLS algorithm is computationally more economical, wherein PS-GMRF texture features are calculated in groups, thereby obviating the large matrix process in the classical LS method.

1) For \( k = 1 \):

\[
\bar{\beta}^{(1)} = \left[ \sum_{s \in S} Q^{(1)}(s) \left( Q^{(1)}(s) \right)^T \right]^{-1} \times \left[ \sum_{s \in S} Q^{(1)}(s) (f(s) - \mu) \right]
\]

\[
\bar{\tau}^{(1)} = \frac{1}{M \times N} \sum_{s \in S} [(f(s) - \mu) - Q^{(1)}(s)^T \bar{\beta}^{(1)}]^2.
\]

2) For \( 1 < k \leq m \):

\[
\bar{\beta}^{(k)} = \left[ \sum_{s \in S} Q^{(k)}(s) \left( Q^{(k)}(s) \right)^T \right]^{-1} \times \left[ \sum_{s \in S} Q^{(k)}(s) (f(s) - \mu + R) \right]
\]

\[
\bar{\tau}^{(k)} = \frac{1}{M \times N} \sum_{s \in S} [(f(s) - \mu - \sum_{i=1}^{k} (Q^{(i)}(s))^T \bar{\beta}^{(i)})]^2
\]

where

\[
R = \sum_{i=1}^{k-1} \left[ \frac{1}{m} (f(s) - \mu) - \left( Q^{(i)}(s) \right)^T \bar{\tau}^{(i)} \right].
\]

From (7)–(10), it is obvious that the parameters in the low-level group are computed independent of those in the higher level groups; however, the parameters in the high-level groups are estimated based on those in the lower level groups. For example, the vector of the parameters in the first group \( \bar{\beta}^{(1)} \) is calculated by (7), not relating to the parameters or neighbors in other groups. However, the vector of parameters in the second group \( \bar{\beta}^{(2)} \) is calculated by (9) with \( k = 4 \), dependent on \( \bar{\beta}^{(1)}, \bar{\beta}^{(2)}, \text{ and } \bar{\beta}^{(3)} \), via the supplement variable \( R \).

\[3\] In the traditional way, model order is one of the important concerns for a successful texture classification. Usually, the more complicated the texture, the higher order model should be used. Kashyap and Chellappa [6] presented a method generated from Bayes procedure for estimation and choice of the method order for the GMRF texture.

B. Texture Feature Selection

A set of texture features having good discriminating power is essential for a texture-image-classification system. The presence of a large number of texture features could lead to poor
classification results if care is not given to the contribution of these features to the interclass separability [16]. In order to reduce data dimensions and improve classification quality, GMRF features extracted from the input image should be processed by feature selection.

The problem of feature subset selection is to select a subset of \( b \) features from a larger set of \( a \) \((b < a)\) features to optimize the value of a criterion over all subsets of the size \( b \). There are \( \binom{a}{b} = (a!/(a-b)!) \) such subsets, for example, \( \binom{11}{3} = (11!/3!8!) = 165 \) and \( \binom{14}{3} = (14!/3!11!) = 364 \). Exhaustive evaluation of all the subsets is computationally prohibitive. To avoid the exhaustive search for an optimal feature subset, direct methods of selecting a suboptimal feature subset instead of searching for an optimal subset have been developed [17], [18], such as sequential forward selection (SFS), sequential backward selection (SBS), plus-l take-away-r (PTA), sequential forward floating selection (SFFS), and sequential backward floating selection (SBFS). SFS (SBS) begins with a feature subset and sequentially adds (removes) features until the termination criterion is met. Both of them suffer from the so-called “nesting effect,” which implies that for SFS the features selected cannot be reselected later. PTA was proposed to prevent the “nesting” effect, repeating this process: Go forward \( l \) stages by adding \( l \) features by SFS and go backward \( r \) stages by deleting \( r \) features by SBS. However, there is no theoretical guidance to determine the appropriate value of \( l \) and \( r \). SFFS and SBFS are the floating version of PTA [19]. SFFS can be understood as sequential forward selection (SFS), sequential backward selection (SBFS), plus-one minus one minus, the floating version of PTA. SFFS can backtrack unlimitedly as long as the backtrack finds a better effect, repeating this process: Go forward \( l \) stages by adding \( l \) features by SBS. However, there is no theoretical guidance to determine the appropriate value of \( l \). SFFS and SBFS are the floating version of PTA. SFFS can backtrack unlimitedly as long as the backtrack finds a better feature subset than the feature subset obtained so far at the same size. Moreover, SBFS is the backward version of SFFS.

In this paper, the SFFS method is adopted in view of its good performance in both the quality of obtained feature subset and computation efficiency. Texture features are evaluated according to their discriminatory power. The criterion used here for feature selection is the Kappa coefficient of testing samples, based on the Gaussian maximum-likelihood classifier [20].

### III. Classification Methodology

Maximum a posterior (MAP) estimation is known to be optimal in the sense that it is a probabilistic relaxation method as a means to enforce the spatial constraints into classification. The fundamental notion of spatial context is that the class attributions of two spatially adjacent pixels are highly related. For example, if \( s \) and \( s' \) are two neighboring pixels and if \( s \) belongs to class \( c \), then there is a high probability that pixel \( s \) also belongs to the same class \( c \). Thus, classification should be carried out based not only on feature values of each pixel but also on the statuses of the corresponding pixels.

An input multidimensional feature image to be classified is described by \( Y = \{y_1, y_2, \ldots, y_w\} \), where \( w \) is the number of pixels in the input image and \( y_i = (y_{i1}, y_{i2}, \ldots, y_{id})^T \) represents a vector of \( d \) features (can be spectral and texture features) for the \( i \)th pixel. Assuming that there are \( C \) predetermined classes in the input feature image, the corresponding classified image can be denoted by \( l = \{l_1, l_2, \ldots, l_w\} \), where \( l_i \) takes on value in \( \{1, 2, \ldots, C\} \), and each pixel is assigned to one of the \( C \) classes according to predefined rules. The MAP estimator is to maximize the following expression [21]:

\[
p(l|y) = \frac{p(y|l)p(l)}{p(y)}
\]

where \( p(l|y) \) is the posteriori probability of \( l \) conditioned on \( y \), \( p(y|l) \) denotes the likelihood distribution of \( y \) conditioned on \( l \), and \( p(y) \) is the probability distribution of \( y \). As \( p(y) \) is normally considered as constant, maximizing (11) is equivalent to maximizing the product of \( p(y|l) \) and \( p(l) \).

In this paper, \( p(y|l) \) is modeled in terms of the Gaussian distribution, while \( p(l) \) is modeled based on the second-order multilevel logistic (MLL) model. When MAP estimation is applied to image classification, the task becomes a combinatorial optimization problem [22]. A deterministic relaxation algorithm, iterated conditional modes (ICM) [23], is used because of its computational efficiency. The ICM solution is obtained by performing the following optimization. Moreover, for the \( i \)th pixel, the class attribute is determined by

\[
l_i = \arg \max_{c \in \{1,2,\ldots,C\}} \{p(y_i|c)p(c|j \in V(i))\}
\]

where \( V(i) \) denotes the eight-connectivity neighbor set of the \( i \)th pixel. The law \( p(y_i|c) \) is given by

\[
p(y_i|c) = \frac{1}{\sqrt{(2\pi)^d|\Sigma_c|}}\exp\left[-\frac{1}{2}(y_i - \mu_c)^T\Sigma_c^{-1}(y_i - \mu_c)\right]
\]

where \( \mu_c \) is the mean vector and \( \Sigma_c \) is the covariance matrix associated with the class \( c \). The chosen a priori term is an MLL model, which yields

\[
p(c|j \in V(i)) = \frac{1}{Z} \exp \left[-\sum_{j \in V(i)} \alpha \delta(c, l_j)\right]
\]

where \( \alpha \) is the model parameter expressing the strength of how an occurrence of class \( c \) for pixel \( i \) is affected by its neighbor class attributes \( \{l_j, j \in V(i)\} \), \( Z \) is a normalizing constant, and \( \delta(c, l_j) \) is an indicator function, if \( c = l_j \), \( \delta(c, l_j) = -1 \), otherwise, \( \delta(c, l_j) = 1 \). In fact, the variation of \( \alpha \) does not greatly influence the results [24], and \( \alpha \) is taken as 1.0 in our experiments. Then, using (13) and (14) and after some manipulations, (12) is equivalent to

\[
l_i = \arg \min_{c \in \{1,2,\ldots,C\}} \left\{\frac{1}{2}(y_i - \mu_c)^T\Sigma_c^{-1}(y_i - \mu_c) + \frac{1}{2}\ln|\Sigma_c| + \sum_{j \in V(i)} \alpha \delta(c, l_j)\right\}.
\]

Equation (15) is to achieve the minimum of the entire energy, which could be considered as the sum of the energy of the
observed field (the first term plus the second term) and the energy of the labeled field (the third term).

The classification method based on the MAP criterion using the ICM update scheme, named as the MAP–ICM algorithm, is implemented by the steps outlined below. The method is described by a flow chart in Fig. 2.

Step 1) Estimate the statistic parameters \((\mu_c, \Sigma_c)\) for each class.

Step 2) Based on \((\mu_c, \Sigma_c)\), estimate an initial classification using the noncontextual pixelwise maximum-likelihood decision rule by minimizing the energy of the observed field of (15).

Step 3) Update \((\mu_c, \Sigma_c)\) by

\[
\mu_c = \frac{1}{n_c} \sum_i y_i I(l_i, c) \quad (16)
\]

and

\[
\Sigma_c = \frac{1}{n_c} \sum_i (y_i - \mu_c)(y_i - \mu_c)^T I(l_i, c) \quad (17)
\]

where \(n_c\) is the number of pixels in class \(c\); \(I(l_i, c) = 1\), if pixel \(i\) belongs to class \(c\) and zero otherwise.

Step 4) For each pixel of the image, perform the local minimization defined by (15) and update the corresponding label in parallel.

Step 5) Go back to Step 3) until the specified determination condition is satisfied.

In this paper, we take the maximum iteration of 80 as the stopping criterion. Note that in Step 1), for the supervised classification, \((\mu_c, \Sigma_c)\) is computed from the training samples of each class, while for the unsupervised classification \((\mu_c, \Sigma_c)\) is obtained by the fuzzy C-means (FCM) algorithm [25]. Supposing a partition matrix \(U = [u_{ci}]\), where \(u_{ci}\) represents the probability for the pixel \(i\) to be in the cluster \(c\), \(\mu_c\) and \(\Sigma_c\) are estimated from the \(u_{ci}\)

\[
\mu_c = \frac{\sum_{i=1}^{w} u_{ci}^2 y_i}{\sum_{i=1}^{w} u_{ci}^2} \quad (18)
\]

IV. CLASSIFYING LAND-COVER TYPES USING QUICKBIRD IMAGERY

In this paper, we consider a QuickBird true-color image with 2.4-m spatial resolution of a suburban area in Beijing, China. Fig. 3 is a mosaic of six different land-cover types selected from the QuickBird image, coded as C1–C6 from left to right and then up to bottom. C1 is the residential districts made of concrete, in which houses are arranged orderly. C2 and C3 are woodlands with different density and have similar spectral characteristics; C2 is sparse and fine in texture, while C3 looks very thick and exhibits large particles in texture. C4, C5, and C6 are three different kinds of crops. From the point of view of reflective spectral values, C5 is different from C4 and C6, and C4 is quite like C6. Moreover, in terms of spatial structures, C5 is coarsest, and C6 is finest. The size of each image chip is 128 × 128.

The classification scheme outlined in Fig. 4 was used for classifying different classes of the QuickBird image. First, band 1 (B-band) of QuickBird data was chosen for computing texture features, because this band has the maximum variability in terms of standard derivation. Texture features are extracted based on a moving window of a fixed size. We attempt to find a window that is not too large in order to limit computation time and avoid the influence of the texture features of adjacent different types; on the other hand, the window should not be too small, as otherwise a robust GMRF estimator would not be obtained [26]. The \(17 \times 17\) pixels were used in the experiments as a result of a tradeoff between the application of too large and too small a window, and the odd number makes sure that the window is centered on the current pixel. Using a fourth-order GMRF model and a program written in the C++ language, a 14-D PS-GMRF feature vector is yielded for each pixel of the input image. All experiments in this paper are performed using the computer with Intel Pentium processor 1.60-GHz and EMS.
texture features are normalized so that the measurements have zero mean and unit standard deviation. Afterwards, SFFS is implemented and the selected PS-GMRF texture features have zero mean and unit standard deviation. The computational time for PS-GMRF texture feature extraction from Fig. 3 is 133 s.

All the extracted texture features are normalized so that the measurements have zero mean and unit standard deviation. Afterwards, SFFS is implemented and the selected PS-GMRF texture features are $\hat{\beta}(0, 1)$, $\hat{\beta}(1, -1)$, and $\hat{\nu}$, extracted from the QuickBird texture mosaic. The average profile of the three selected PS-GMRF features is shown in Fig. 7(d). In the same way, the histograms of the selected CGMRF texture features: $\tilde{\beta}(0, 1)$, $\tilde{\beta}(1, -1)$, and $\tilde{\nu}$ are plotted in Fig. 7(a)–(c) and the average profile in Fig. 8(d). A good feature should have minimal intraclass distance and maximum interclass distance. The histograms in Fig. 7 are approximately detached, while those in Fig. 8 almost coincide for the texture types. Obviously, PS-GMRF features are more discriminatory than CGMRF features. Moreover, the transformed divergence (TD) measures between class pairs of Fig. 3 are given in Table I(a) and (b), using the selected PS-GMRF and CGMRF features, respectively. TD values vary from 0.0 to 2.0 corresponding to complete overlap and ideal separation between two classes. For the selected PS-GMRF features, the overall average of TD values among classes was 1.99, whereas it dropped to 1.31 for the selected CGMRF features.

Next, in order to evaluate the contribution of the integration of the spectral and texture features in the discrimination of the six land-cover types, the spectral data and the texture data were, respectively, used alone to classify the QuickBird texture mosaic. The comparison results are listed for five combinations in Table II: 1) using spectral features alone; 2) using the selected CGMRF texture features alone; 3) combining the spectral and selected CGMRF features; 4) using the selected PS-GMRF texture features alone; and 5) combing the spectral and selected PS-GMRF features. From this table, two valuable observations available, the performance of the two types of texture features can be exactly evaluated in a quantitative manner. Overall accuracy and Kappa coefficient, which are proven to be general and robust for any classification algorithm (pixel-based classification methods, neighbor/region-based classification methods [27], [28], or object-based classification methods), are acted as criteria to summarize classification results [29].
can be made. One is that PS-GMRF texture features provide better discrimination than CGMRF texture features. The Kappa coefficient increased when using the selected PS-GMRF features (0.8474) compared to classification using the selected CGMRF features (0.3026). The other is that the integrated utilization of spectral and texture features can obtain a more satisfactory result than either spectral or texture features alone. The Kappa coefficient was 0.6214 using spectral data alone, 0.8474 using the selected PS-GMRF texture data alone, and 0.9281 using the combination of spectral and selected PS-GMRF texture data.

V. RESIDENTIAL AREA EXTRACTION USING LOW-ORDER VARIANCE

Residential areas are characterized by low correlation and high variance, which cannot be accurately detected by spectral data. It was found that one of the PS-GMRF texture features, the lowest order variance $\nu^{(1)}$, is effective for residential area extraction. The following introduces the experiments using the IKONOS and SPOT-5 images. These images were provided by the Space Imaging Corporation and the Satellite Imaging Corporation.

A. Experiment on IKONOS Imagery

Fig. 9(a) shows a true-color image collected by the IKONOS satellite and is situated at the Brazil/Paraguay border. This image consists of three scenes of 350 lines, each 400 pixels in length. When using the spectral features alone, the classified image was obtained with the MAP–ICM algorithm, as shown in Fig. 9(b). The initialization was processed by means of FCM, in which the number of clusters was fixed at five beforehand. From
Fig. 9. (a) Initial IKONOS image. (b) Classified result of (a) using the spectral data.

Fig. 10. PS-GMRF texture features extracted from Fig. 9(a). (a) Feature image of $\nu(1)$, (b) $\nu(2)$, (c) $\nu(3)$, and (d) $\nu(4)$.

In order to reduce the blurring border effect, inherent in texture analysis and which introduces important errors in the transition areas between different texture units [12], a classification procedure was designed for such satellite images as follows. First, residential areas are detected using texture features; afterward, other objects are classified using spectral features. Based on the fourth-order GMRF model, PS-GMRF texture features were computed from the one band with the largest spectral variability. Fig. 10 displays the results of the extraction of four PS-GMRF features: $\nu(1)$, $\nu(2)$, $\nu(3)$, and $\nu(4)$. It is easy to see that the lowest order variance $\nu(1)$ has a larger capacity for discriminating between residential areas and other land-cover types. Two clusters, one a residential area and the other not, were detected in the lowest order variance space using FCM. Sequentially, a mask was generated to eliminate residential areas so that other land-cover types would be classified accurately, and not interfered with the spectrally heterogeneous residential areas. The MAP-ICM classification algorithm under the mask was performed on the input data in Fig. 9(a), followed by combining the resident map, as shown in Fig. 11(a).

Fig. 11. (a) Overlap image between the detected residential areas (in white lines) and the original IKONOS image. (b) Final classified result of Fig. 9(a) after combining detected residential areas.

In order to prove the effectiveness of the lowest order variance for residential-area detection in a quantitative manner, comparisons of discriminative abilities of fourth-order variances ($\nu(1)$, $\nu(2)$, $\nu(3)$, and $\nu(4)$) are provided based on the classification result of Fig. 11(b). For the sake of simplicity, five different types of land covers are coded. The residential area labeled in white in Fig. 11(b) is denoted as D1, and other land covers are numbered by D2–D5 given in Table III. Just similar to Section IV, 2000 samples for each kind of land covers were selected according to the final classified image in Fig. 11(b), and histograms were plotted to analyze distributions of the fourth-order variances corresponding to each land-cover type, shown in Fig. 12. All of PS-GMRF texture features were normalized to have zero mean and unit standard deviation. Fig. 12 contains four graphs (a)–(d), corresponding to $\nu(1)$, $\nu(2)$, $\nu(3)$, and $\nu(4)$, respectively. Each graph has five curves of histogram statistics and each curve represents the distribution of one feature with regard to one type of land covers. In Fig. 12, the graph (a) has the largest gaps between D1 (residential areas) and D2–D5 (other types) and no overlapped area between them. It is not difficult to find that the first-order variance $\nu(1)$ has the higher ability to discriminate between residential areas and other land covers than those higher order variances, such as $\nu(2) - \nu(4)$.

<table>
<thead>
<tr>
<th>Grey Labeling Types of Land Covers</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Residential area)</td>
<td>D1</td>
</tr>
<tr>
<td>(Water)</td>
<td>D2</td>
</tr>
<tr>
<td>(Forest)</td>
<td>D3</td>
</tr>
<tr>
<td>(Beach)</td>
<td>D4</td>
</tr>
<tr>
<td>(Grass)</td>
<td>D5</td>
</tr>
</tbody>
</table>

Fig. 11. (a) Final classified result of Fig. 9(a) after combining detected residential areas.

TABLE III
Coded About Five Types of Land Covers in Fig. 11(b)
B. Experiment on SPOT-5 Imagery

Subsequent experiments were made on a SPOT-5 image. Fig. 13(a) is part of the SPOT-5 satellite image of Reno, NV, acquired on September 14, 2003, and the image size is $400 \times 400$. Using the spectral features, the MAP–ICM classification with FCM was carried out to detect five clusters. The result is shown in Fig. 13(b), and residential areas are not separated accurately from other land-cover types. Residential areas with high intraclass spectral variance would be detected exactly when texture features are considered. The band with the largest spectral variability was selected from Fig. 13(a) and PS-GMRF texture features computed with the fourth-order GMRF model. Fig. 14 shows the PS-GMRF variances: $\nu^{(1)}$, $\nu^{(2)}$, $\nu^{(3)}$, and $\nu^{(4)}$. In Fig. 14(a), the values are high for pixels inside the residential areas, which are characterized by high lowest order variance, while the values are low for pixels in other land covers, such as fields and water areas, which are characterized by small lowest order variance. To avoid the adverse effect of blurring at the boundaries, residential areas were first extracted according to the lowest order variance. White lines outline the detected residential areas in Fig. 15(a). After masking the residential areas, the MAP–ICM procedure is processed in Fig. 13(a), in which FCM is used to initialize four clustering centers. By integrating the above two results, the final classified image was achieved, as shown in Fig. 15(b). The proposed classification approach improves the classification accuracy and visual interpretation, compared to classification with purely spectral features.

Subsequently, capabilities of residential-area extraction are presented for different order variances to rededuce the above conclusion that the lowest order variance for the detection of residential areas is more effective than the higher order variances, based on the classification result. Five land-cover types are numbered E1–E5, as listed in Table IV. E1 represents residential areas, E2–E5 as the other land covers. In the
Corresponding to each order variance were given in Fig. 16, where Fig. 16(a)–(d) show the statistical calculation, sequentially the distribution curves of the lowest order variance according to the classified image Fig. 15(b), were used for the same way, for each type of land cover, 2000 samples selected, except for one land-cover types of Fig. 15(b).

The selected PS-GMRF texture features, in combination with spectral features, are used instead of PS-GMRF features for comparison. The Kappa coefficient increased from 0.6214 to 0.6621 for an addition of the selected CGMRF texture features. It is obvious that PS-GMRF features are superior to CGMRF features.

The authors would like to thank the Space Imaging Corporation for providing the IKONOS image and the Satellite Imaging Corporation for providing the SPOT-5 image.

### ACKNOWLEDGMENT

The authors would like to thank the Space Imaging Corporation for providing the IKONOS image and the Satellite Imaging Corporation for providing the SPOT-5 image.

### REFERENCES


Y. Hu and T. J. Dennis, “Simulated annealing and iterated conditional
J. Besag, “On the statistical analysis of dirty pictures,”
S. Krishnamachari and R. Chellappa, “Multiresolution Gauss–Markov
A. Clausi and H. Deng, “Operational segmentation and classification
X. P. Jia and J. A. Richards, “Efficient maximum likelihood classification
M. Kudo and J. Sklansky, “Comparison of algorithms that select features
X. P. Jia and J. A. Richards, “Efficient maximum likelihood classification
B. Tso and R. C. Olsen, “Scene classification using combined spectral,
Yindi Zhao received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2006.
Liangpei Zhang (M’06) received the B.S. degree in physics from Hunan Normal University, ChangSha, China, in 1982, the M.S. degree in optics from the Xi’an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi’an, in 1988 and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.
From 1997 to 2000, he was a Professor with the School of the Land Sciences, Wuhan University. In August 2000, he joined the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, as a Professor and Head of the Remote Sensing Section. He has published more than 120 technical papers. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.
Dr. Zhang has served as the Cochair of the International Society for Optical Engineers Series Conferences on Multispectral Image Processing and Pattern Recognition (MIPPR) and the Conference on Asia Remote Sensing in 1999, as an Editor of the MIPPR’01, MIPPR’05, and Geoinformatics Symposiums, as an Associate Editor of Geo-spatial Information Science Journal, as a member of the Chinese National Committee for the International Geosphere-Biosphere Programme, and as an Executive Member for China Society of Image and Graphics.
Pingxiang Li (M’06) received the B.S., M.S., and Ph.D. degrees in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1986, 1994, and 2003, respectively.
Since 2002, he has been a Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. His research interests include photogrammetry and SAR image processing.
Bo Huang received the Ph.D. degree in geographic information system from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 1997.
He is currently an Associate Professor with the Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. His current research interests include spatial optimization, spatial statistics for change analysis, and image processing. He has authored and coauthored over 30 papers in refereed international journals and developed several GIS, spatial analysis, and optimization software packages.