Pergamon

Neural Networks, Vol. 10, No. 3, pp. 561–572, 1997

Printed in Great Britain
0893–6080/97 $17.00+.00

PII: S0893-6080(96)0092-5

CONTRIBUTED ARTICLE

Multispectral Image Classification Using Gabor Filters and Stochastic Relaxation Neural Network

P. P. RAGHU AND B. YEGNANARAYANA

Indian Institute of Technology

(*Received* 6 *February* 1996; *accepted* 30 *May* 1996)

Abstract—*In this article, we propose a supervised classification scheme for multispectral image data based on the spectral as well as textural features. A filter bank consisting of Gabor wavelets is used to extract the features from the multispectral imagery. The classification model consists of three random processes, namely, feature formation, partition and label competition. The feature formation process models the multispectral texture features from the Gabor filter bank as a multivariate Gaussian distribution. The partition process and the label competition process represent a set of label constraints. These constraints are represented on a Hopfield neural network model, and a stochastic relaxation strategy is used to evolve a global minimum energy state of the network, corresponding to the maximum a posteriori (MAP) probability. The performance of the scheme is demonstrated on a variety of multispectral multipolar images obtained from SIR-C/X-SAR.* © *1997 Elsevier Science Ltd. All Rights Reserved.*

Keywords—Multispectral image classification, Texture-based analysis, Constraint satisfaction neural networks, Gabor filters, Stochastic relaxation strategy, Remote sensing, Synthetic aperture radar images.

1. INTRODUCTION

Classification of remotely sensed images is essential for efficient interpretation of the imagery. While remote sensing procedure significantly reduces the complexity of the task of actual measurements as compared to the in situ methods, the extraction of relevant information from the image data is still an involved procedure.

An important feature of remote sensing is the multispectral nature of the image data. A number of studies demonstrated the use of multispectral information for image classification (Jensen, 1986). In our studies, we have used the synthetic aperture radar (SAR) images which are characterized by both multispectral and multipolar behavior. In the discussions hereafter, the term multispectral images includes both multispectral and multipolar images from SAR.

Many land types in a remotely sensed imagery appear as a distribution of spatial patterns of tonal variation consisting of repetition or quasi-repetition of some fundamental image elements, controlled by parameters such as surface topology, color and morphology. Thus, each land cover type is characterized by an underlying texture-like pattern which may be different from those of other land types. Several studies have reported the usefulness of textures in remote sensing, for example, in the discrimination of crops using radar imagery (Berger, 1970), in the terrain classification using aerial photography (Weszka et al., 1976), and in the analysis and classification of sea-ice types using SAR data (Holmes et al., 1984).

In computer-based analysis of multispectral images, classifications based only on multispectral pixel information are usually inefficient to discriminate different land-cover classes. In such situations, one has to use both spectral and textural information to provide better accuracy in classification. One attempt which uses spectral as well as textural features was made by Shih and Schowengerdt (1983) for the classification of arid geomorphic surfaces from Landsat MSS data. In a study, Bischof et al. (1992) showed that the incorporation of textural information improves the accuracy of multispectral classification of Landsat TM imagery.

Recently, the resurgence of interest in artificial neural networks has brought into focus the use of such models for different tasks in pattern recognition and image processing. But conventional neural network models may not perform satisfactorily on the images obtained from remote sensing. This is due to the specific nature and

Requests for reprints should be sent to Professor B. Yegnanarayana, Department of Computer Science and Engineering, Indian Institute of Technology, Madras-600 036, India; Tel: 91-044-235-1365 ext. 3505; Fax: 91-044-235-0501; e-mail: yegna@iitm.ernet.in.

characteristics of the remotely sensed images. The networks should be capable of dealing with partial and noisy information and should efficiently take care of spatial and spectral characteristics.

In this paper, we describe a neural network approach for the classification of multi-spectral images based on the underlying textural and spectral information. Each spectral image is filtered using the Gabor wavelets and the resulting feature vectors are concatenated to obtain the feature vector representing each pixel in the multispectral imagery. The feature extraction using Gabor filters overcomes some of the issues in texture analysis of remotely sensed images. The joint probability of the extracted features is modeled as multivariate Gaussian distribution using a feature formation process, the parameters of which are estimated from the features of pixels in the training sites. Two other random processes, namely partition process and label competition process, define a set of label constraints. The classification model used in our work is defined by the a posteriori probability derived from these random processes and is expressed as a Gibb's distribution. A Hopfield network is used to represent the corresponding Gibb's energy function as a set of constraints on each pixel. A stochastic relaxation mechanism for the network finds the global minimum of the Gibb's energy. The equilibrium state of the network yields the optimal classification of the given imagery.

We have used a similar concept for segmenting the natural textures using the Hopfield neural network model with a deterministic relaxation strategy (Raghu & Yegnanarayana, 1996). The work reported in this paper is an extension of the earlier work to classify multispectral textures from remote sensing. In contrast with the earlier method, the present work makes use of a multivariate Gaussian distribution as the feature model in order to capture the interband and intraband correlations of the Gabor wavelet features from the multispectral imagery. It also uses a stochastic relaxation strategy in order to overcome the local minima problems in the Hopfield network.

Some of the symbols used in the paper are described here. Consider the textured multispectral imagery T_Ω designated by a domain $\Omega = \{(i,j), 0 \leq i < I, 0 \leq j < J\}$ of pixel positions. Let the imagery comprise of B bands of images, the b^{th} band image being denoted by $T_{b,\Omega}$. Let $Y_{b,s}$, $s \in \Omega$ be the random variable corresponding to the gray level value of the pixel s in the b^{th} band image and can take any integer value in the range $\{0,...,255\}$. Assume the gray value of a pixel s to be $y_{b,s}$ as an instantiation of $Y_{b,s}$. A set of M' dimensional feature vectors $\{g_{b,s} \in R^{M'}, \forall s \in \Omega\}$ generated by Gabor filtering is used to characterize the image $T_{b,\Omega}$. Each pixel s in the entire imagery T_Ω is represented by an M dimensional vector $g_s = \biguplus_b g_{b,s}$ where $M = BM'$ and $\uplus$ denotes vector concatenation. Each g_s can be considered to be the realization of an M dimensional random process G_s.

The proposed classification scheme assumes that the imagery T_Ω consists of K different number of textures, so that each pixel s can take any texture label 0 to $K-1$. The corresponding texture classes are denoted by $C_0,..., C_{K-1}$. Also, let Ω_k, a subset of Ω, be the training site for the class C_k. The Gabor features of the training site of a given class are used to estimate the model parameters for that class. We use the notation L_s to denote the random variable describing the texture label of the pixel s.

The paper is organized as follows. The next section deals with some issues in the texture analysis of images from remote sensing. The discussion leads to the necessity of specific feature extraction methods for such images. In Section 3, the multiresolution feature extraction mechanism based on Gabor filters is described. This includes a discussion on the advantages of Gabor filters which make them suitable for texture analysis in remote sensing. The classification model comprising of feature formation process, partition process and label competition process is presented in Section 4 and the corresponding neural network representation is described in Section 5. Finally, the performance of the proposed classification scheme for classifying different images from SIR-C/X-SAR is discussed in Section 6.

2. ISSUES IN TEXTURE ANALYSIS OF IMAGES FROM REMOTE SENSING

Remotely sensed images have a number of characteristics which make the subsequent image processing and analysis unique. As discussed earlier, an important one is the multispectral nature of imagery. Efficient use of multispectral information for the interpretation of imagery is still an active research problem.

Depending upon the spatial resolution of the remote sensor and the fineness of underlying textures, an image can have both textured regions and intensity regions. The image analysis methods should be able to process the images irrespective of whether the image contains regions which are textured, nontextured or both.

The images from remote sensing are usually noisy and sparse due to several atmospheric conditions and sensor characteristics. This necessitates the features considered for classification of imagery to be immune to noise.

In an image containing a number of textures, texture elements may have different sizes and shapes, making it difficult to determine a priori the resolution for texture analysis and hence to define the size of analysis window for feature extraction. The problem of determining appropriate resolution for texture analysis is referred to as the *cell unit problem* (Wechsler, 1980). Conventional monoresolution methods fail to capture features of varying sizes.

In an image from remote sensing, the features may be spatially varying within a textured region. Such textures are called nonstationary textures (Reed & Wechsler, 1988). Land cover types in remotely sensed images

may have textural patterns which are gradually varying within their region, still remaining in the same land type. Conventional feature extraction methods assume the stationarity of texture within the analysis window and hence may not be useful for the textures from remote sensing. We need to characterize the textures in spatial as well as spatial-frequency domains simultaneously by using joint spatial/spatial frequency (s/sf) representations (Jacobson & Wechsler, 1988).

In the present work, we specifically address the issues of resolution, nonstationarity and noisy data with reference to the images from remote sensing.

3. MULTIRESOLUTION FEATURE EXTRACTION USING GABOR FILTERS

One efficient way to deal with the cell unit problem is by reorganizing the image into a number of subsampled approximations of it at different resolutions. This is called multiresolution analysis (Rosenfeld, 1982). A method to achieve this is the wavelet representation (Mallat, 1989). This scheme analyses the coarse image details first and gradually increases the resolution to analyse the finer details. The local variations in the orientation and frequency of texture elements lead to the nonstationary behavior of remotely sensed textures. To capture this, we need a joint spatial/spatial-frequency representation which is orientation-sensitive. In this section, we describe the use of Gabor filters as bases for the wavelet decomposition of images to extract texture features. This resembles the mechanism of multichannel representation of the retinal images in the biological visual system (Daugman, 1980).

A 2 D Gabor filter is an oriented complex sinusoidal grating modulated by 2 D Gaussian function. The expression for the 2 D Gabor wavelet is given as,

$$f(x,y,a,\omega,\theta,\sigma)=e^{-\frac{a^2}{2\sigma^2}(x^2+y^2)+ja\omega(x\cos\theta+y\sin\theta)} \quad (1)$$

In this expression, θ and ω are the orientation and radial frequency of the Gabor filter. The bandwidth σ of 2 D Gaussian gives the spatial resolution of the filter. The parameter a is the wavelet scale factor, which is chosen as $a = 2^{\gamma}$, where γ is an integer. A set of filters $\{f(x,y,a,\omega,\theta,\sigma)\}_{a,\theta}$ constitute the Gabor wavelet family consisting of a number of scaled and rotated versions of the mother wavelet.

The use of Gabor wavelets provides varying support in spatial as well as spatial frequency domains in order to detect and localize optimally the texture features at different scales. A number of studies have been made on the use of Gabor filters for the analysis of textures (Bovik et al., 1990). The following are the advantages of Gabor filters in analyzing textures from remote sensing:

1. 2 D Gabor filters represent an image both in spatial and spatial frequency domains optimally by achieving the theoretical lower bound of joint uncertainty (Daugman, 1985). This makes it possible to extract characteristic features from nonstationary textures.
2. The orientation selectivity of Gabor filters discriminates textures possessing different orientations.
3. The arrangement of Gabor filters as wavelet bases provides multiresolution representation and makes it possible to capture details of different sizes.
4. Gabor filters are able to extract multispectral texture features in presence of noise (Raghu et al., 1993). This is useful especially in analysing SAR images which are characterized by speckle noise.

For the b^{th} band image $T_{b,\Omega}$, the feature value at position (x, y) is given by

$$g_{b,\lambda}(x,y)=|T_{b,\Omega}(x,y) * f_{\lambda}(x,y)|^2 \quad (2)$$

for a filter $f_{\lambda}(x,y)$ with a given parameter set $\lambda = (a,\omega,\theta,\sigma)$. Here, * denotes 2 D convolution. Assume that the Gabor wavelet family contains M' Gabor filters to filter each band of the imagery. For a given pixel point (x,y) in the b^{th} band image, the M' dimensional vector $g_{b,(x,y)} = [g_{b,\lambda}(x,y)]_{\lambda}$ for all M' values of λ constitutes the feature vector to characterize the pixel (x,y) of the image. As mentioned earlier, $g_{(x,y)} = \biguplus_b g_{b,(x,y)}$ forms the feature vector to represent a pixel (x,y) of the multispectral imagery.

4. CLASSIFICATION MODEL

The classification model is a hierarchical one consisting of three different random processes, namely feature formation process, partition process and label competition process.

4.1. Feature Formation Process

The feature formation process describes the probability of assigning a value $g_s \in R^M$ to the random process G_s of pixel s given the model parameters of each texture class. Let us define this conditional probability as the multivariate Gaussian distribution given by

$$P(G_s=g_s|L_s=k)=\frac{1}{\sqrt{(2\pi)^M|\Sigma_k|}}e^{\frac{1}{2}(g_s-\mu_k)^{\top}\Sigma_k^{-1}(g_s-\mu_k)} \quad (3)$$

where $\top$ is the transpose operation. This can also be written as

$$P(G_s=g_s|L_s=k) = e^{-\frac{1}{2}\left[(g_s-\mu_k)^{\top}\Sigma_k^{-1}(g_s-\mu_k)+\ln\left((2\pi)^M|\Sigma_k|\right)\right]} \quad (4)$$

The model parameters μ_k and Σ_k for each class C_k are estimated from the training site Ω_k of that class as,

$$\text{Mean } \mu_k = \frac{1}{S_k} \sum_{s \in \Omega_k} g_s \tag{5}$$

$$\text{Covariance matrix } \Sigma_k = \begin{bmatrix} v_{11k} & v_{12k} & \cdots & v_{1Mk} \\ v_{21k} & v_{22k} & \cdots & v_{2Mk} \\ \cdots & \cdots & \cdots & \cdots \\ v_{M1k} & v_{M2k} & \cdots & v_{MMk} \end{bmatrix} \tag{6}$$

where each component v_{ijk} is the covariance of each element g_s(i) and g_s(j) of g_s and is estimated as,

$$v_{ijk} = \frac{1}{S_k} \sum_{s \in \Omega_k} [g_s(i) - \mu_k(i)][g_s(j) - \mu_k(j)] \tag{7}$$

S_k is the cardinality of set Ω_k. The importance of covariance matrix is that it characterizes the interband correlation as well as the intraband correlation of the Gabor filtered multispectral imagery.

4.2. Partition Process

The label of any pixel in an image depends on the labels of the pixels in its neighborhood. The partition process $P(L_s|L_r, \forall r \in N_s^p)$ describes probability of the label of each pixel s given the labels of the pixels in a uniform p^{th} order neighborhood N_P^s of s (Raghu & Yegnanarayana, 1996). This can be modeled as a p^{th} order Markov random field model defined by,

$$P(L_s|L_r, \forall r \in N_s^p) = \frac{e^{\beta \sum_{\forall r \in N_s^p} \delta(L_s - L_r)}}{Z_p} \tag{8}$$

where β is a positive constant, $\delta(.)$ is the Kronecker delta function and Z_p is a normalization constant.

4.3. Label Competition Process

The label competition process attempts to reduce the probability of having another label when the pixel is already labeled (Raghu & Yegnanarayana, 1996). It is defined by the conditional probability,

$$P(L_s = k|L_s = l)_{\forall l} = \frac{e^{-\alpha \sum_{\forall l} \bar{\delta}(k-1)}}{Z_c} \tag{9}$$

where k is the new label for the pixel s, and l stands for any label already assigned to s. In this expression, α is a positive constant, Z_c is a normalization constant $\bar{\delta}$ is the inverse of the Kronecker delta function given by

$$\bar{\delta}(l) = \begin{cases} 0, & \text{if } l = 0 \\ 1, & \text{otherwise} \end{cases} \tag{10}$$

4.4. Formulation of a Posteriori Probability

The classification model is defined by the a posteriori probability, $P(L_s = k|G_s, L_r, \forall r \in N_s^p, L_s = l)_{\forall l}$. This describes the labeling L_s of the pixel s given the feature measurement of s, the labels of the neighborhood pixels and the possible labels previously assigned to s. This can be written using Bayes theorem as

$$P(L_s = k|G_s, L_r, \forall r \in N_s^p, L_s = l)_{\forall l} = \frac{P(G_s|L_s = k)P(L_s = k|L_r, \forall r \in N_s^p)P(L_s = k|L_s = l)_{\forall l}}{P(G_s)P(L_s = k)} \tag{11}$$

We express the a posteriori probability as a Gibbs distribution given by

$$P(L_s = k|G_s, L_r, \forall r \in N_s^p, L_s = l)_{\forall l} = \frac{e^{-E(L_s = k|G_s, L_r, \forall r \in N_s^p, L_s = l)_{\forall l}}}{Z} \tag{12}$$

where, by substituting eqns (4), (8) and (9) in eqn (11), we get the Gibb's energy as

$$E(L_s = k|G_s, L_r, \forall r \in N_s^p, L_s = l)_{\forall l} = \frac{1}{2}[(g_s - \mu_k)^T \Sigma_k^{-1}(g_s - \mu_k) + \ln((2\pi)^M |\Sigma_k|)] - \sum_{\forall r \in N_s^p} \beta\delta(k - L_r) + \sum_{\forall l} \alpha\bar{\delta}(k - l) \tag{13}$$

and the normalization constant Z is $Z_p Z_c P(G_s = g_s) P(L_s = k)$. Assuming any pixel s has constant a priori probabilities of having a feature vector g_s and any label k, Z can be considered as constant irrespective of s and k.

The energy function in eqn (13) summed over all pixels and all possible labels will give the total energy E^{total} of the classification model.

$$E^{total} = \sum_{s,k} \left[\frac{1}{2}[(g_s - \mu_k)^T \Sigma_k^{-1}(g_s - \mu_k) + \ln((2\pi)^M |\Sigma_k|)] - \sum_{\forall r \in N_s^p} \beta\delta(k - L_r) + \sum_{\forall l} \alpha\bar{\delta}(k - l) \right] \tag{14}$$

Estimation of a state configuration L_s for all pixels s which minimizes the Gibb's energy in eqn (14) will yield an optimal classification of the multispectral imagery. This is equivalent to maximizing the a posteriori probability given in eqn (11). A neural network

model based on the Hopfield network is proposed for this purpose. A stochastic optimization procedure based on simulated annealing is used to find a global (or near-global) minimum energy state.

5. CLASSIFICATION USING A STOCHASTIC RELAXATION NEURAL NETWORK

The energy function in eqn (14) can be represented on a Hopfield network with nodes arranged in a 3 dimensional lattice of size $I \times J \times K$. For any node (i,j,k), the pixel position and the corresponding label index are denoted by (i,j) and k, respectively. The connection weight between any two nodes (i,j,k) and (i_1,j_1,k_1) is denoted by $W_{i,j,k;i_1,j_1,k_1}$ and is assumed to be symmetric, i.e. $W_{i,j,k;i_1,j_1,k_1} = W_{i_1,j_1,k_1,i,j,k}$. Each node (i,j,k) in the network can have an external bias $B_{i,j,k}$. Assume $\Lambda_{i,j,k} \in \{0,1\}$ to be the output of node (i,j,k). The set $\kappa = \{\Lambda_{i,j,k}, \forall i,j,k\}$ is called the state of network. At any instant, $\Lambda_{i,j,k} = 1$ indicates that the pixel (i,j) has a label k at that instant. We use the notation $\Lambda_{i,j,k}(n)$ to denote the output of the node at n^{th} iteration of relaxation algorithm.

Comparing eqn (14) with the energy function of the Hopfield network (Hopfield, 1982), the bias $B_{i,j,k}$ and the weight $Wi,j,k;i_1,j_1,k_1$ can be determined in a manner similar to Raghu and Yegnanarayana (1996) and Chellappa et al. (1992) as

$$B_{i,j,k} = -\frac{1}{2}\left[(g_s - \mu_k)^{\top} \Sigma_k^{-1} (g_s - \mu_k) + \ln\left((2\pi)^M |\Sigma_k|\right)\right] \tag{15}$$

$$W_{i,j,k;i_1,j_1,k_1} = \begin{cases} 2\beta, & \text{if } (i_1,j_1) \in N^p_{(i,j)} \quad \text{and} \quad k = k_1 \\ -2\alpha, & \text{if } (i_1,j_1) = (i,j) \quad \text{and} \quad k \neq k_1 \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

Initially, each component in the state of the network is randomized with a value in $\{0,1\}$.

The state corresponding to the maximum a posteriori probability is attained by using the simulated annealing procedure (Kirkpatrick et al., 1983). The probability of the network being in a state κ is assumed obeying the *Boltzmann–Gibbs distribution*, i.e. $P(\kappa) \propto e^{-\frac{1}{T}E(\kappa)}$ where $E(\kappa)$ is the energy of network at the state κ. T is pseudo-temperature which controls updation of states during the relaxation procedure. The probability of transition from a state κ_1 to κ_2 is derived as

$$P(\kappa_1 \rightarrow \kappa_2) = \frac{1}{1 + e^{\frac{1}{T}\Delta E_{\kappa_1 \rightarrow \kappa_2}}} \tag{17}$$

$\Delta E_{\kappa_1 \rightarrow \kappa_2} = E(\kappa_2) - E(\kappa_1)$ is the change of energy when the state changes from κ_1 to κ_2.

The temperature term T introduces uncertainty in evolving a stable state in the network. A high temperature means a high noise to the system which helps it to escape from local minima and also prevents it from settling anywhere. At lower temperatures, convergence becomes slower, making finer discrimination among different energy states. At zero temperature, the network behaves exactly like the Hopfield network. The simulated annealing procedure starts the state updation at high temperature and gradually decreases the temperature to a minimum value. This ensures that the system has the best chance of ending in a state corresponding to the lowest energy minimum with respect to a given network, no matter what state it was started in. It is because this method, unlike the deterministic state updation in Hopfield network (Hopfield, 1982), allows with non-zero probability to temporarily change the state towards a higher energy state, thus leading to an escape from a local minimum.

The simulated annealing procedure sequentially updates each node selected at random. Let κ_1 and κ_2 be the states of the network before and after updating the output $\Lambda_{i,j,k}(n)$ of a node (i,j,k) at n^{th} iteration. The $\Delta E_{\kappa_1 \rightarrow \kappa_2}$ can be derived from the Hopfield energy function as Kung (1993), chapter 2,

$$\Delta E_{\kappa_1 \rightarrow \kappa_2} = -U_{i,j,k}(n)\Delta\Lambda_{i,j,k}(n+1) \tag{18}$$

where, $U_{i,j,k}(n)$ is the net input of node (i,j,k) at n^{th} iteration given by

$$U_{i,j,k}(n) = \sum_{i_1,j_1,k_1} W_{i,j,k;i_1,j_1,k_1} \Lambda_{i_1,j_1,k_1}(n) + B_{i,j,k} \tag{19}$$

and $\Delta\,\Lambda_{i,j,k}(n+1) = \Lambda_{i,j,k}(n+1) - \Lambda_{i,j,k}(n)$, the change in the output of that node. The probability of changing the output $\Lambda_{i,j,k}(n)$ of node (i,j,k) is,

$$P(\kappa_1 \rightarrow \kappa_2) = \frac{1}{1 + e^{\frac{-1}{T}U_{i,j,k}(n)\Delta\Lambda_{i,j,k}(n+1)}} \tag{20}$$

The output of the selected node is updated according to the probability of state transition given in eqn (20). The temperature is initialized to a high temperature T_0. At each temperature, the network state is updated until thermal equilibrium occurs at that temperature. Subsequently, the temperature is reduced using an annealing schedule and the state update is continued in that temperature. For our experiments, the exponential decay,

$$T = T_0 e^{-n/\tau} \tag{21}$$

is used as the annealing schedule to get a new temperature. Here, τ is the decay time constant. The annealing and subsequent state updates are continued until the temperature is reduced to a small value. At that instant, the classification is interpreted from the label of each pixel

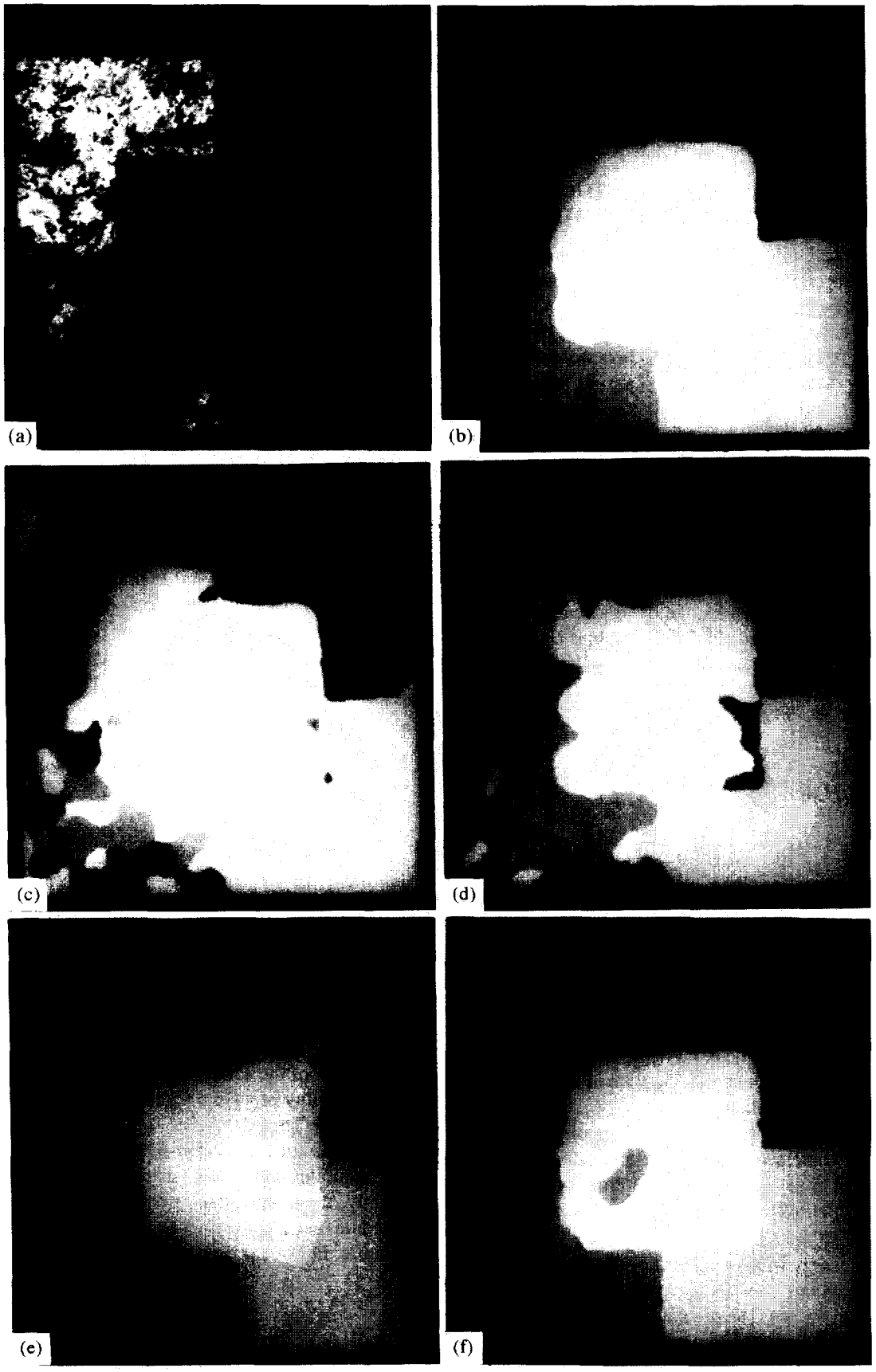

FIGURE 1. Classification of textures. (a) Image containing texture tiles (five classes). (b) Classification using the proposed scheme. (c) Classification using the red channel image. (d) Classification using the green channel image. (e) Classification using the blue channel image. (f) Classification using the raw imagery.

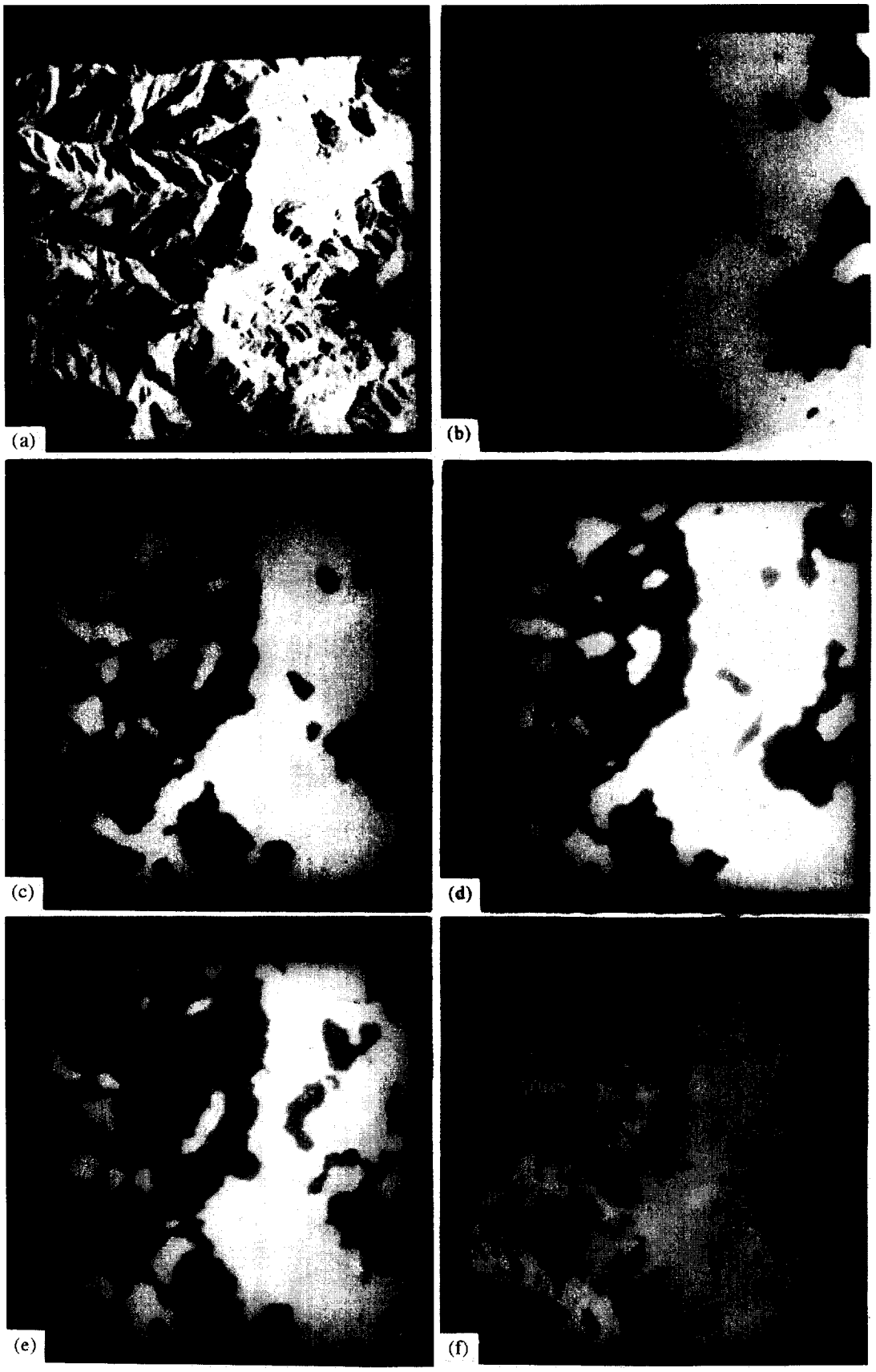

FIGURE 2. Classification of textures. (a) Image of a region from Mount Pinatubo, Philippines (two classes). (b) Classification using the proposed scheme. (c) Classification using the red channel image. (d) Classification using the green channel image. (e) Classification using the blue channel image. (f) Classification using the raw imagery.

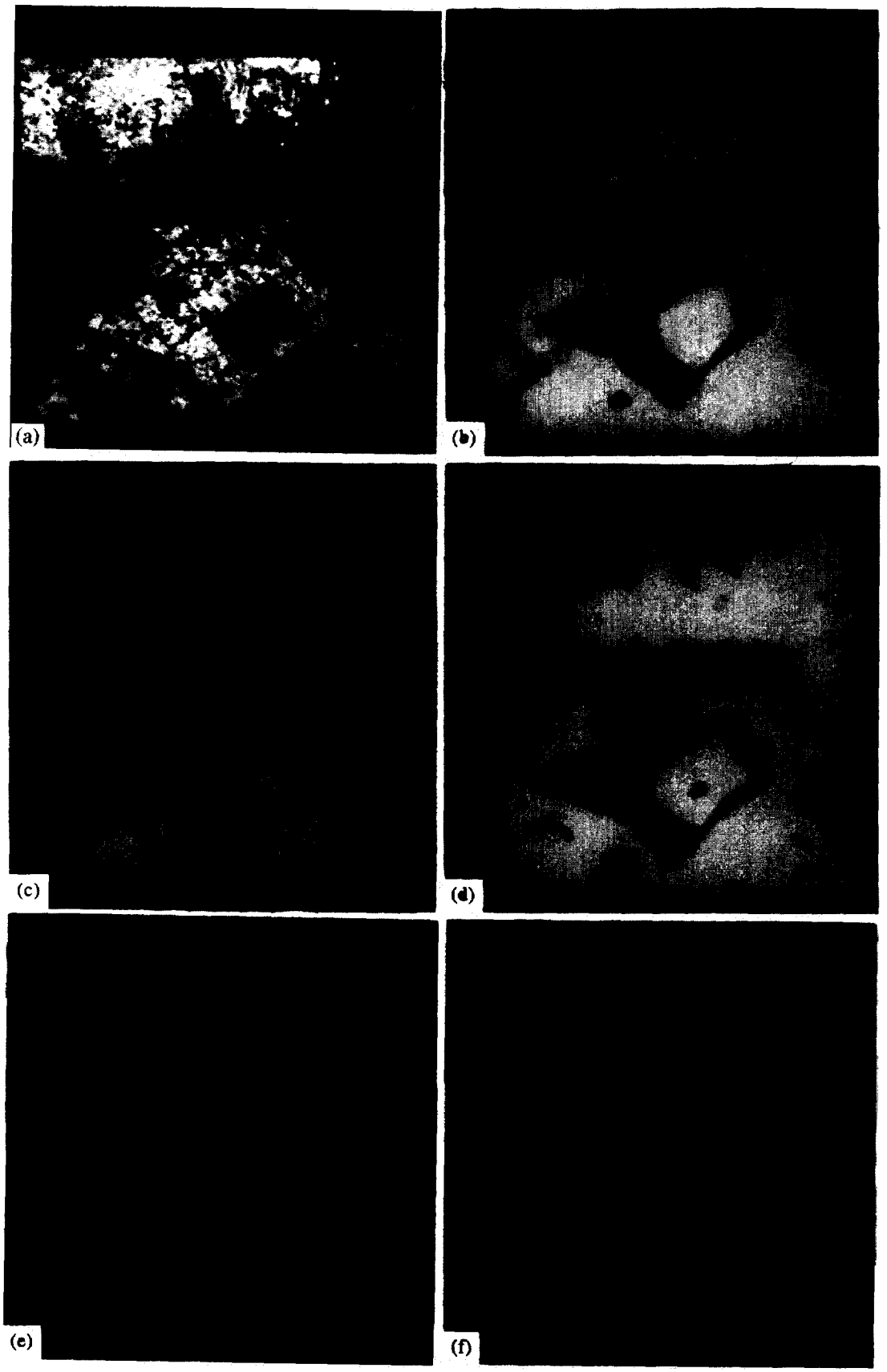

FIGURE 3. Classification of textures. (a) Image of a region from Flevoland, Netherlands (four classes). (b) Classification using the proposed scheme. (c) Classification using the red channel image. (d) Classification using the green channel image. (e) Classification using the blue channel image. (f) Classification using the raw imagery.

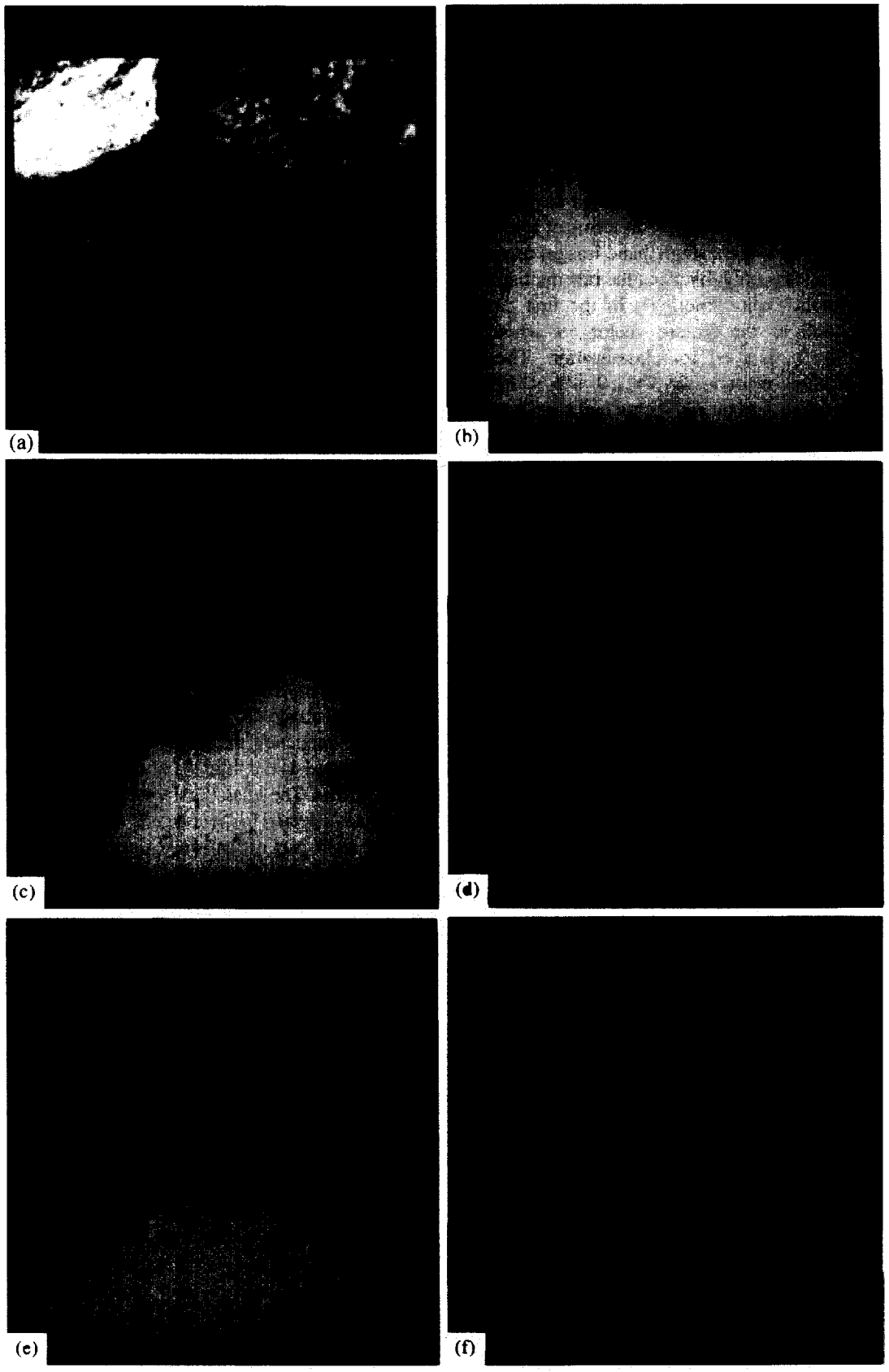

FIGURE 4. Classification of textures: (a) Image of a region from the site of the lost city of Ubar, South Oman (four classes). (b) Classification using the proposed scheme. (c) Classification using the red channel image. (d) Classification using the green channel image. (e) Classification using the blue channel image. (f) Classification using the raw imagery.

given as

$$\text{Label of the pixel } (i,j) = \max_k U_{i,j,k} \qquad (22)$$

where $U_{i,j,k}$ is the net input of the node (i,j,k) corresponding to the final state of the network.

To compare the performance of the proposed method which classifies the multispectral images using textural and spectral information, we have conducted the experiments to classify the given imagery in terms of textural information only (by using a single band image) as well as spectral information only (by using the raw multispectral image without Gabor filtering). In the first case, $\{g_{b,s}, \forall s\}$ is used as the feature vector set when a single band, say b, is used for classification. The B dimensional feature vector $y_s = \{y_{b,s}, 0 \leq b < B\}$ is used as feature for each pixel s in the second case, which is a pixel-based multispectral classification approach.

6. EXPERIMENTAL RESULTS

The Spaceborne Imaging Radar-C and X-Synthetic Aperture Radar (SIR-C/X-SAR) (Jordan et al., 1991) of NASA/Jet Propulsion Laboratory acquires images in three microwave bands, namely, L-band (24 cm), C-band (6 cm) and X-band (3 cm) with four linear polarization states, namely, horizontal transmit horizontal receive (HH), horizontal transmit vertical receive (HV), vertical transmit horizontal receive (VH) and vertical transmit vertical receive (VV).

A number of images from SIR-C/X-SAR were considered, to study the performance of the proposed classification scheme. The first image (Fig. 1a) is created by concatenating different multispectral texture tiles. The other images (Fig. 2a, Fig. 3a and Fig. 4a) are characterized undefined texture boundaries and unknown texture models. The results in each case show the original image, the classification result by using the proposed classification scheme, the results when each single band data is used for classification (without multispectral information), and the result when classification is performed using the raw image (without textural features).

Fig. 1a shows the image of a three channel image with size 256 $\times$ 256 pixels consisting of five multispectral texture tiles. To extract texture features from each of the spectral images, a filter bank consisting of eight Gabor filters was used. The filters were derived from a mother wavelet of $\sigma = 25$ and $\omega = 0.2$ by using two scales ($\gamma = 3$ and 4, with wavelengths of 1.25π and 0.625π pixels/cycle) and four orientations ($\theta = 0$, 45, 90, and 135 degrees). The concatenation results in a 24 dimensional feature vector to characterize each pixel in the three-band imagery. Fig. 1b shows the classification using the proposed scheme. The results of using only single band information are given in Fig. 1c, d and e, with the red, green and blue channels, respectively, of a false color image. The result when the image is classified by using only multispectral information is shown in Fig. 1f. The classification result in Fig. 1b shows the performance of the proposed scheme on the image containing multispectral texture tiles. It indicates the ability of the scheme in correctly classifying the class regions and in locating the class boundaries with a fairly good accuracy. In contrast, the schemes which use either textural information alone (Fig. 1c–e) or multispectral information alone (Fig. 1f) are not able to capture the behavior of each class. There are many misclassified regions in these results.

The next data used for experiments consists of the imagery of a region extracted from the area around Mount Pinatubo in Philippines. The image shown in Fig. 2a gives the regions affected with ash deposit during volcano eruptions. For supervised classification, we have identified two texture classes in the imagery – the ash area and the underlying rock regions. Fig. 2b shows multispectral textural classification using the proposed scheme. The Gabor filters used for this imagery are same as those of the previous experiment. Single band texture classification results are shown in Fig. 2c–e with red, green and blue channels. Fig. 2f shows the multispectral classification without textural features. It can be seen from the result given in Fig. 2b that the proposed scheme which uses both multispectral and textural features clearly brings out the two different classes present in the image.

Three-frequency imagery of Flevoland, The Netherlands, is used as the third set of data for the classification experiments. The image is shown in Fig. 3a. The number of classes selected in the image is four, with forest region (upper right part of the image), urban area (upper left), water body (middle area), and bare soil in the agricultural land (lower part of the image). We used two Gabor filters per band (with $\sigma = 1.5, \omega = 3.2, \gamma = 0$ and $\theta = 0, 90$ degrees), constituting 6 dimensional feature vector for each pixel in the imagery. The classification using proposed scheme is provided in Fig. 3b. Classification results using red, green and blue channels of a false color image are shown in Fig. 3c, d and e, respectively, and multispectral pixel classification is shown in Fig. 3f. The results shown in Fig. 3c–f break the imagery into too many smaller regions, which indicates the inadequacy of multispectral or textural information alone in describing the land cover types. In contrast with this, the result in Fig. 3b, which combines this information, gives a fairly acceptable classification of the given imagery.

Fig. 4a shows the radar image of the region around the site of the lost city of Ubar in Southern Oman. The imagery is assumed to have four classes. We detail these classes as follows. The region in the lower part of image is a region of large sand dunes. The area in the upper right part of the image is rough lime stone rock and the upper left region is a dry stream bed (wadi). The sandy area which lies inbetween the lime stone rocks and the sand dunes has similar spectral properties as

TABLE 1
Different Parameters used for the Classification Experiments

Image in Figure	α	β	T_0	τ	p
1a	1.0	0.060	10000	3.5	6
2a	1.0	0.220	10000	6.0	2
3a	1.0	0.075	10000	3.0	4
4a	1.0	0.080	10000	5.0	9

that of sand dunes, but different textural properties. The transition from this sandy area to sand dune region is not clearly defined and this makes the segmentation difficult for these two classes. For this imagery, we used 8 Gabor filters per band, having the same parameters as those for the first imagery. Fig. 4b shows the classification based on multispectral-textural information. Fig. 4c–f shows the results using red channel, green channel, blue channel and raw imagery respectively of a false color image. There are two notable aspects of these results. The proposed scheme, the result of which is shown in Fig. 4b, has classified the sand dune region in a remarkably good manner irrespective of the nonstationary textural patterns present in that region. Moreover, the algorithm has been able to clearly locate the boundary between sandy area and the region of sand dunes. These are the major achievements in the result, while the classification using either the multispectral or the textural information is not able to achieve this, and the corresponding results contain many spurious misclassified regions. Different parameters used for the classification experiments of each imagery are given in Table 1.

7. CONCLUSION

In this paper, we have presented a supervised multispectral classification approach for classifying and segmenting textured multispectral images from remote sensing. Specific issues of processing textures in remote sensed images were addressed by using the feature extraction based on Gabor filters. The dependence of feature vectors on the class labels was modeled as a multivariate Gaussian distribution. The constraints regarding pixel labels were formulated using partition and label competition processes. These constraints were represented on a Hopfield neural network model, and the simulated annealing incorporated in this network was used to find the maximum a posteriori estimate of the optimal image classification. A number of textured images from SIR-C/X-SAR were used to demonstrate the performance of the proposed method. It can be inferred from the classification results that the proposed scheme, which combines textural as well as the multispectral features, provides much better performance compared to the classification based on either textural features or multispectral features.

REFERENCES

Berger, D.H. (1970). Texture as a discriminant of crops on radar imagery. *IEEE Transactions on Geoscience and Electronics*, **GE-8 (4),** 344–348.

Bischof, H., Schneider, W., & Pinz, A.J. (1992). Multispectral classifications classification of Landsat-images using neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, **30 (4),** 482–490.

Bovik, A.C., Clark, M., & Geisler, W.S. (1990). Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12 (1),** 55–73.

Chellappa, R., Manjunath, B. S., & Simchony, T. (1992). Texture segmentation with neural networks. In Bart Kosko (Ed.), *Neural Networks for Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.

Daugman, J.G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, **20,** 847–856.

Daugman, J.G. (1985). Uncertainty relation for resolution in space, spatial-frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of Optical Society of America (A)*, **2 (7),** 1160–1169.

Holmes, Q.A., Nüesch, D.R., & Shuchman, R.A. (1984). Textural analysis and real time classification of sea-ice types using digital SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **GE-22 (2),** 113–120.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the USA*, **79,** 2554–2558.

Jacobson, L.D., & Wechsler, H. (1988). Joint spatial/spatial frequency representations. *Signal Processing*, **14 (1),** 37–68.

Jensen, J. R. (1986). *Introductory Digital Image Processing: A Remote Sensing Perspective*. Englewood Cliffs, NJ: Prentice-Hall.

Jordan, R.L., Huneycutt, B.L., & Werner, M. (1991). The SIR-C/X-SAR synthetic aperture radar system. *Proceedings of the IEEE*, **79 (6),** 827–838.

Kirkpatrick, S., Gelatt, C.D. Jr., & Vecchi, M.P. (1983). Optimization by simulated annealing. *Science*, **220,** 671–680.

Kung, S. Y. (1993). *Digital Neural Networks*. Englewood Cliffs, NJ: Prentice-Hall.

Mallat, S.G. (1989). Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoustics Speech and Signal Processing*, **37 (12),** 2091–2110.

Raghu, P. P, Chouhan, H. M., & Yegnanarayana, B. (1993). Multispectral texture classification using neural network. In *Proceedings of the National Conference on Neural Networks* (pp. 1–10). Anna University, Madras, India.

Raghu P. P., & Yegnanarayana, B. (1996). Segmentation of Gabor filtered textures using deterministic relaxation, *IEEE Transactions on Image Processing*, **5**(12).

Reed, T., & Wechsler, H. (1988). Tracking of nonstationarities for texture fields. *Signal Processing*, **14,** 95–102.

Rosenfeld, A. (1982). *Multiresolution Image Processing and Analysis*. New York: Springer-Verlag.

Shih, E.H.H., & Schowengerdt, R.A. (1983). Classification of arid geomorphic surfaces using Landsat spectral and textural features. *Photogrammetric Engineering and Remote Sensing*, **49 (3),** 337–347.

Wechsler, H. (1980). Texture analysis – a survey. *Signal Processing*, **2**, 271–282.

Weszka, J.S., Dyer, C.R., & Rosenfeld, A. (1976). A comparative study of texture measures for terrain classification. *IEEE Transactions on System, Man and Cybernetics*, **SMC-6**, 269–285

NOMENCLATURE

a	Wavelet scaling factor chosen as $a = 2^{\gamma}$ where γ is an integer
B	Number of bands in the multispectral imagery
$B_{i,j,k}$	Bias of the node (i,j,k) in the 3 D Hopfield network
C_k	k^{th} texture class, $k \in \{0,...,K-1\}$
E_f, E_p, E_c	Energy functions of feature formation, partition and label competition processes
$f(x,y,a,\omega,\theta,\sigma)$	Gabor wavelet with radial frequency ω, bandwidth σ, orientation θ and wavelet scaling factor a
G_s	M D random process which denotes the Gabor feature vector at pixel $s \in \Omega$
g_s	Instantiation of G_s
I,J	Size of each spectral image T_b,Ω
K	Number of texture classes
L_s	Random variable denoting label of pixel s; $L_s \in \{0,...,K-1\}$
M	Dimension of the Gabor feature space of the multispectral imagery; $M = BM'$
M'	Number of Gabor functions used to filter each spectral band image
MAP	Maximum a posteriori
M D	M dimensional
N_L^p	Set of displacement vectors corresponding to a p_{th} order symmetric neighborhood of image pixels
p	Order of the neighborhood used in the partition process
R^M	M D space of real numbers
SAR	Synthetic Aperture Radar
SIR-C/X-SAR	Spaceborne Imaging Radar-C and X-Synthetic Aperture Radar
$\top$	Transpose operation of a matrix
T_0	Initial temperature in the simulated annealing
T_b,Ω	b^{th} band image in the multispectral imagery
T_Ω	Multispectral imagery containing B number of spectral bands
$U_{i,j,k}(n)$	Net input of the node (i,j,k) in the 3 D Hopfield network at n^{th} iteration of the relaxation
$W_{i,j,k;i_1,j_1,k_1}$	Connection weight from a node (i,j,k) to (i_1,j_1,k_1) of the 3 D Hopfield network
$Y_{b,s}$	Random variable corresponding to the intensity value of a pixel s in the b^{th} band image $T_b\Omega$; $Y_{b,s} \in \{0,...,255\}$
$y_{b,s}$	Instantiation of $Y_{b,s}$
-2α, 2β	Connection weights of the 3 D Hopfield network
γ	An integer used in the wavelet scaling factor a
θ	Orientation of the Gabor filter
$\Lambda_{i,j,k}(n)$	Output of the node (i,j,k) in the 3 D Hopfield network at n^{th} iteration of the relaxation
σ	Bandwidth of the Gabor filter
τ	Decay time constant used in the annealing schedule
Ω	Domain of pixel positions; $\Omega = \{(i,j), 0 \leq i < I, 0 \leq j < J\}$
Ω_k	Training site for the class C_k; $\Omega_k \subset \Omega$
ω	Radial frequency of Gabor filter