The Use of Fully Polarimetric Information for the Fuzzy Neural Classification of SAR Images

Chia-Tang Chen, *Member, IEEE*, Kun-Shan Chen, *Senior Member, IEEE*, and Jong-Sen Lee, *Fellow, IEEE*

***Abstract*—This paper presents a method, based on a fuzzy neural network, that uses fully polarimetric information for terrain and land-use classification of synthetic aperture radar (SAR) image. The proposed approach makes use of statistical properties of polarimetric data, and takes advantage of a fuzzy neural network. A distance measure, based on a complex Wishart distribution, is applied using the fuzzy c-means clustering algorithm, and the clustering result is then incorporated into the neural network. Instead of preselecting the polarization channels to form a feature vector, all elements of the polarimetric covariance matrix serve as the target feature vector as inputs to the neural network. It is thus expected that the neural network will include fully polarimetric backscattering information for image classification. With the generalization, adaptation, and other capabilities of the neural network, information contained in the covariance matrix, such as the amplitude, the phase difference, the degree of polarization, etc., can be fully explored. A test image, acquired by the Jet Propulsion Laboratory Airborne SAR (AIRSAR) system, is used to demonstrate the advantages of the proposed method. It is shown that the proposed approach can greatly enhance the adaptability and the flexibility giving fully polarimetric SAR for terrain cover classification. The integration of fuzzy c-means (FCM) and fast generalization dynamic learning neural network (DLNN) capabilities makes the proposed algorithm an attractive and alternative method for polarimetric SAR classification.**

***Index Terms*—Fuzzy c-means, neural classification, polarimetric synthetic aperture radar, speckle filtering.**

I. Introduction

A NUMBER of classification algorithms have been proposed for polarimetric synthetic aperture radar (SAR) images. As stated in [1], there are basically three approaches: 1) algorithms based on image processing techniques, 2) algorithms based on a statistical model, and 3) algorithms based on the scattering mechanism of electromagnetic waves. These approaches can be either supervised or unsupervised. Supervised and unsupervised algorithms complement each other; each has its own advantages and disadvantages, depending on applications, and the availability of ground truth. For example, Lee *et al.* [20] used the statistical properties of a fully polarimetric SAR to perform a supervised classification based on a complex Wishart distribution. However, the mixed-pixel problem was not considered, even though fully polarimetric features were used. Du and Lee [11], [22] integrated a complex Wishart distribution and fuzzy c-means (FCM) clustering in an unsupervised technique. Neural networks has the capability of adopting varied types of feature vector while retaining highly flexible in terms of their structure and learning algorithms [2], [3], [5]–[10], [12]. A neural classifier can be either supervised or unsupervised, depending on the learning process and the architecture. Several neural networks for polarimetric SAR classification have been proposed [2], [3], [5], [6], [12]. Feature vectors, derived from the principal components, intensities, or phases between polarizations, have been used in these approaches. However, only partial polarimetric information has been utilized, leaving out the fully polarimetric information. Hence, full advantage has not yet been taken of polarimetric data when neural networks are used as the classifier. This does not necessarily mean that partial polarimetric data is not sufficient for the applications cited, but it does means that the full utilization of polarimetric data does seem to be necessary.

Thus, in this paper, a supervised fuzzy neural classifier for fully polarimetric SAR, which incorporates polarimetric statistics into a fuzzy dynamic learning neural network (FDL) [12], is developed. When calculating the dissimilarity of the FCM (the first part of the FDL), a statistical distance measure, based on a complex Wishart distribution, is employed. In the data preprocessing phase, it is found that the presence of the speckle noise reduces the classification accuracy. A polarimetric SAR filter, as proposed by Lee *et al.* [4], is applied to reduce the speckle noise level, while preserving the polarimetric properties.

The organization of the remainder of the paper is as follows. Section II gives the fundamentals of the polarimetric SAR image, followed by a review of the statistical properties [13]. A statistical fuzzy neural classifier is then developed in Section III. Section IV describes the experimental test results. Finally, the conclusion is given in Section V.

II. Polarimetric SAR Images

A. Statistical Properties

The polarimetric SAR records the scattering matrix $\mathbf{S}$, which represents the relationship between the transmitted and received electric field [14]

$$\begin{bmatrix} E_h \\ E_v \end{bmatrix}^{\text{rec}} = \frac{e^{ikr}}{kr}\,\mathbf{S}\begin{bmatrix} E_h \\ E_v \end{bmatrix}^{\text{trans}} = \frac{e^{ikr}}{kr}\begin{bmatrix} S_{\text{hh}} & S_{\text{hv}} \\ S_{\text{vh}} & S_{\text{vv}} \end{bmatrix}\begin{bmatrix} E_h \\ E_v \end{bmatrix}^{\text{trans}} \quad (1)$$

Manuscript received October 3, 2001; revised March 21, 2003. This work was supported by the National Council of Agriculture and the National Science Council of Taiwan under Grant NSC90-2212-E-008-037.

C.-T. Chen is with the Department of Information Management, Hsing-Wu College, Taipei, 244 Taiwan, R.O.C.

K.-S. Chen is with the Center for Space and Remote Sensing Research, National Central University, Chung-Li, Taiwan 320, R.O.C.

J.-S. Lee is with the Remote Sensing Division, Naval Research Laboratory, Washington, DC 20375-5351 USA.

Digital Object Identifier 10.1109/TGRS.2003.813494

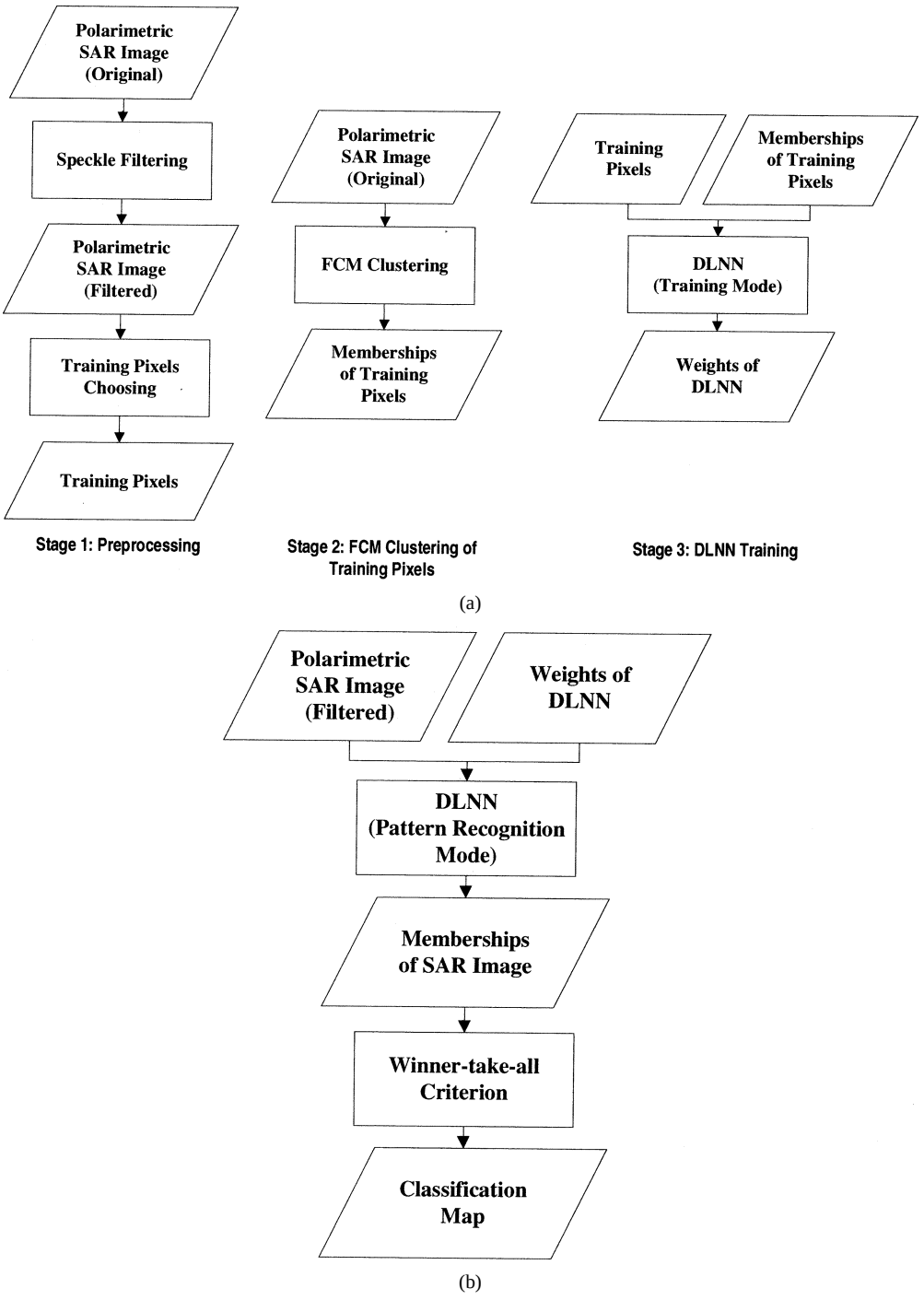


Fig. 1. Flowchart of the proposed statistical fuzzy neural classifier. (a) Phase 1: preprocessing and training. (b) Phase 2: classification.

where the subscripts of $\mathbf{S}$ denote the polarized states; h and v are horizontal and vertical polarization; k is incident wavenumber; r is the range from the radar antenna to the target's center; $trans$ and rec represent the transmitted and the received waves. When the complex scattering matrix $\mathbf{S}$ is available, interactions between the radar waves and the target medium can be fully described in terms of complete polarization responses. For simplicity, a complex vector $\mathbf{k}$ can be defined by the elements of $\mathbf{S}$ as

$$\mathbf{k} = [\, S_{\rm hh} \quad S_{\rm hv} \quad S_{\rm vh} \quad S_{\rm vv} \,]^T \tag{2}$$

where T denotes the matrix transposition. For backscattering from a reciprocal medium, $S_{\rm hv} = S_{\rm vh}$. In case of backscattering, the scattering vector $\mathbf{k}$ can be rewritten as

$$\kappa = [\, S_{\rm hh} \quad \sqrt{2}\, S_{\rm hv} \quad S_{\rm vv} \,]^T. \tag{3}$$

The polarimetric covariance matrix is defined as

$$\mathbf{C} = \mathbf{k}\mathbf{k}^{*T} = \begin{bmatrix} S_{\rm hh}S_{\rm hh}^* & \sqrt{2}\, S_{\rm hh}S_{\rm hv}^* & S_{\rm hh}S_{\rm vv}^* \\ \sqrt{2}\, S_{\rm hv}S_{\rm hh}^* & 2S_{\rm hv}S_{\rm hv}^* & \sqrt{2}\, S_{\rm hv}S_{\rm vv}^* \\ S_{\rm vv}S_{\rm hh}^* & \sqrt{2}\, S_{\rm vv}S_{\rm hv}^* & S_{\rm vv}S_{\rm vv}^* \end{bmatrix} \tag{4}$$

where $*$ denotes the complex conjugate.

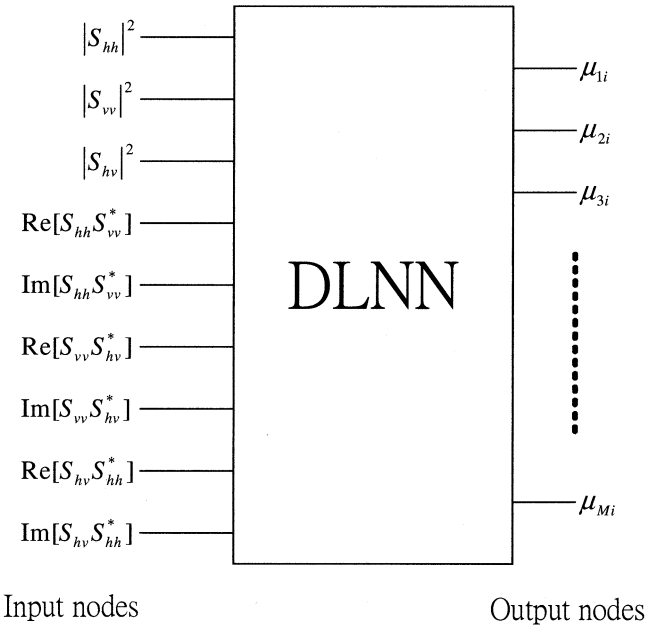


Fig. 2. Configuration of the DLNN for adopting the complete polarimetric SAR data inputs. The input nodes of the DLNN are the elements of feature vector, where $|S_{\mathrm{hh}}|^2$, $|S_{\mathrm{hv}}|^2$, $|S_{\mathrm{vv}}|^2$ are the diagonal terms in polarimetric covariance matrix and $S_{\mathrm{hh}}S_{\mathrm{vv}}^*$, $S_{\mathrm{vv}}S_{\mathrm{hv}}^*$, $S_{\mathrm{hv}}S_{\mathrm{hh}}^*$ are the off-diagonal terms in polarimetric covariance matrix. The polarimetric covariance matrix is Hermitian so that the six off-diagonal terms can be reduced to three terms. Also, all the off-diagonal terms are complex values and are expressed as real(Re[]) and imagery(Im[]) parts. The values at the output nodes of neural network are the membership of a given pixel to certain classes.

SAR is a coherent imager, and suffers from speckle noise, which degrades image quality. Thus, SAR images are usually multilook processed by averaging several neighboring one-look pixels. For a polarimetric SAR, an average of several one-look covariance matrices is required [15]

$$\mathbf{z} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{C}_i = \begin{bmatrix} \langle S_{\mathrm{hh}}S_{\mathrm{hh}}^*\rangle & \langle\sqrt{2}\,S_{\mathrm{hh}}S_{\mathrm{hv}}^*\rangle & \langle S_{\mathrm{hh}}S_{\mathrm{vv}}^*\rangle \\ \langle\sqrt{2}\,S_{\mathrm{hv}}S_{\mathrm{hh}}^*\rangle & \langle 2S_{\mathrm{hv}}S_{\mathrm{hv}}^*\rangle & \langle\sqrt{2}\,S_{\mathrm{hv}}S_{\mathrm{vv}}^*\rangle \\ \langle S_{\mathrm{vv}}S_{\mathrm{hh}}^*\rangle & \langle\sqrt{2}\,S_{\mathrm{vv}}S_{\mathrm{hv}}^*\rangle & \langle S_{\mathrm{vv}}S_{\mathrm{vv}}^*\rangle \end{bmatrix}_i \quad (5)$$

where C_i is the one-look covariance matrix of the ith pixel, and N is the number of looks. The $\mathbf{z}$ statistics have a complex Wishart distribution [16]

$$P_n(\mathbf{z}) = \frac{n^{qn}|\mathbf{z}|^{n-q}\exp\left[-n\mathrm{Tr}(\boldsymbol{\omega}^{-1}\mathbf{z})\right]}{Q(n,q)|\boldsymbol{\omega}|^n}$$
$$Q(n,q) = \pi^{(1/2)q(q-1)}\Gamma(n)\cdots\Gamma(n-q+1) \quad (6)$$

where $\boldsymbol{\omega} = E[\mathbf{z}]$; q is the dimension of $\mathbf{k}$; $q = 4$ for the bistatic case; and $q = 3$ for the backscatter case.

B. Polarimetric Speckle Filtering

A pixel in an SAR image is formed by the vector sum of echoes from all scatterers within the illuminated cell. All scatterers within a resolution cell may not be electromagnetically homogeneous and their geolocations in relation to the radar are random. A random walk phenomenon occurs that causes speckle noise. SAR image interpretation suffers from this speckle contamination. Thus, the first step for the classification of an SAR image is to reduce the effect of speckle. For properly filtering fully polarimetric SAR data, adaptive filters [17], [18] may be inadequate, because cross-talk between polarization channels is introduced, which destroys the polarimetric information, (e.g., phase differences between different polarized channels). Lee *et al.* [4] developed a polarimetric filter to avoid cross-talk, while preserving the edge sharpness and, most importantly, the polarimetric properties. In [4], it is demonstrated that the classification of fully polarimetric SAR data can be significantly improved by using a new polarimetric filter. Although the neural classifier proposed in this study is free from the statistical distribution of the data assumption, i.e., that no *a priori* information, the preservation of the polarimetric properties is essential in terms of using a fully polarimetric image. In what follows, we shall apply a polarimetric filter to some test data as a preprocessing step. The filtering effects and their impact will be illustrated.

III. Supervised Fuzzy Neural Classifier for Polarimetric SAR

Chen *et al.* [6] developed a dynamic learning neural network (DLNN) for the neural-based classification of polarimetric SAR and have applied it to classify multipl-polarization SAR images. The DLNN is a supervised neural classifier having the advantage of faster learning than the back-propagation neural network. An FDL consisting of FCM clustering and a DLNN was proposed [12]. The first part of the FDL is FCM clustering applied to the training pixels only, not all the pixels of an image. The second part of the FDL is the DLNN which is trained with the feature vectors and the FCM produces memberships of training pixels. Thus, a well-trained DLNN produces memberships of all the pixels classifying them into a specific class that has the maximum membership.

The proposed classifier is a statistical fuzzy neural classifier based on FCM clustering, Wishart distance measure, and DLNN. The feature vector used for classification consists of all elements of the polarimetric covariance matrix, as shown in (5). To accommodate a fuzzy classifier to fully polarimetric features, a distance based on complex Wishart distribution is employed. We discuss this in more detail as follows. Fig. 1 illustrates the flow charts of the whole procedure, which is divided into a training phase and a classification phase. Stage 1 of the training phase is a preprocessing step for noise removal using the polarimetric SAR speckle filter [4]. Next, the training pixels are selected from the filtered image according to a ground survey and an updated base map (in 1/5000 scale or better). Stage 2 of the training phase is FCM clustering on the training pixels. Note that only the training pixels are included in the FCM clustering (not all pixels). In Stage 3, the membership of the training pixels is the desired DLNN outputs. The DLNN is trained with feature vectors (see Fig. 2) to match the desired outputs. Once the DLNN's training is completed, the classification phase is in order. Subsequently, the classification map can be generated by the winner-take-all criterion.

Fig. 3. False-color SAR image of the San Francisco Bay Area. (a) Original. (b) Filtered (with window size of 5×5) (R: |HV|, G: |HH|, B: |VV|).

Fig. 4. False-color SAR images of the Au-Ku plantation. (a) Original. (b) Filtered (with window size of 7×7) (R: |HH $-$ VV|, G: |HV|, B: |HH $+$ VV|).

Now, it is worthwhile to mention the computational load of the proposed approach. If we assume that there are N pixels in the image to be classified into M classes, with N_t training pixels, then by applying FCM clustering alone, the iteration process proceeds with an $N \times M$ matrix [membership function (8)]. By using FDL, the FCM clustering is iterated with the $N_t \times M$ matrix. Usually N is much greater than N_t, so DLNN is performed in pattern recognition mode pixel by pixel. It can be seen that much less computation time is required for FDL than for FCM.

A. Fuzzy Clustering

In the fuzzy c-means algorithm, the position vector of a class or cluster center $\boldsymbol{\omega}_c$ is assumed to be the feature vector average of all the patterns in the class (or cluster) c. In this algorithm, a membership function μ_{ci} is introduced to weight the distance measure and to define the problem of finding the fuzzy c-partitions with a fuzzy index m, where $m \in [1, \infty)$, i.e., to adjust the position vector of the cluster center $\boldsymbol{\omega}_c$ by minimizing the cost function

$$J(\mathbf{U}, \boldsymbol{\Omega}) = \sum_{i=1}^{N_t} \sum_{c=1}^{M} (\mu_{ci})^m \, d^2(\mathbf{z}_i, \boldsymbol{\omega}_c) \tag{7}$$

where $\mathbf{U}$ is a fuzzy c-partition of $\mathbf{Z}$, $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{N_t})$ is a set of N_t training vectors; $\boldsymbol{\Omega} = (\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_M)$ is the cluster center of $\mathbf{Z}$; $d(\mathbf{z}_i, \boldsymbol{\omega}_c)$ is the distance measure between the ith training pattern and the cluster c.

A fuzzy c-partition of input vectors $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{N_t})$ specifies the degree of membership of each vector to each of c classes

$$\mathbf{U} = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1N_t} \\ \mu_{21} & \mu_{22} & & \mu_{2N_t} \\ \vdots & & \ddots & \vdots \\ \mu_{M1} & \mu_{M2} & \cdots & \mu_{MN_t} \end{bmatrix} \tag{8}$$

where μ_{ci} is the degree of membership of pattern $\mathbf{z}_i$ in cluster c. The membership μ_{ci} must satisfy the following conditions [19]:

$$\mu_{ci} \in [0, 1], \qquad 1 \le c \le M, 1 \le i \le N_t$$

$$\sum_{c=1}^{M} \mu_{ci} = 1, \qquad 1 \le i \le N_t$$

$$0 < \sum_{i=1}^{N_t} \mu_{ci} < N_t, \qquad 1 \le c \le M.$$

In [19], the distance measure, or the measure of dissimilarity, under normal distribution, is defined as $d^2(\mathbf{z}_i, \boldsymbol{\omega}_c) = \|\mathbf{z}_i - \boldsymbol{\omega}_c\|^2$, where $\|\cdot\|$ is the inner product norm. In order to include full polarimetric information, the distance measure may be defined by [20]

$$d_w(\mathbf{z}_i, \boldsymbol{\omega}_c) = n\left[\ln|\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i)\right] - \ln[P(\boldsymbol{\omega}_c)] \tag{9}$$

where $\boldsymbol{\omega}_c$ is the cluster center of the covariance matrix for class c, and Tr is the trace of a matrix. Equation (9) represents the distance based on a complex Wishart distribution, with $P(\boldsymbol{\omega}_c)$ as the *a priori* probability of class c. Without *a priori* information, an equal probability is assumed for each class. The distance measure reduces to

$$d_w(\mathbf{z}_i, \boldsymbol{\omega}_c) = \ln|\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i). \tag{10}$$

Differentiating the cost function with respect to $\boldsymbol{\omega}_c$ and μ_{ci}, we follow a similar procedure to [19], under the constraint $\sum_{c=1}^{M} \mu_{ci} = 1, \ \forall i$. After some mathematical manipulations, we obtain

$$\boldsymbol{\omega}_c = \frac{\sum_{i=1}^{N_t} (\mu_{ci})^m \mathbf{z}_i}{\sum_{i=1}^{N_t} (\mu_{ci})^m}, \qquad \forall c \tag{11}$$

and

$$\mu_{ci} = \frac{1}{\sum_{l=1}^{M} \left(\frac{d_w(\mathbf{z}_i, \boldsymbol{\omega}_c)}{d_w(\mathbf{z}_i, \boldsymbol{\omega}_l)}\right)^{2/(m-1)}}, \qquad \forall c, i. \tag{12}$$

The derivations of (11) and (12) can be found in the Appendix. The fuzzy c-means algorithm iteratively solves the above two equations until $\|\mathbf{U}^{(p)} - \mathbf{U}^{(p-1)}\| \le \varepsilon$, where $\|\cdot\|$ denotes the norm of the matrix, p is the count of the iteration, and ε is the criteria for stopping the iteration. The fuzzy c-means algorithm is comprised of the following steps.

TABLE I
CLASSIFICATION CONFIGURATIONS BASED ON THE DIFFERENT FEATURE VECTORS AND THE DISTANCE MEASURES

Setup	Feature vector	Scale	Distance equation in cost function (7)
A	HH, HV, VV	linear	Euclidean
B	HH, HV, VV	logarithmic	Euclidean
C	Covariance matrix	linear	Euclidean
D	Covariance matrix	linear	(9)

1. For a given number of classes, M, $2 \le M \le N_t$, and a fuzzy index, m, $1 \le m < \infty$, and a stopping criterion, $\varepsilon > 0$, initialize $U^{(o)} \in M_{fc}$ and $p = 1$, where M_{fc} is the fuzzy c-partitioned space. The fuzzy index m, in theory, can be any number larger than unity. As m becomes larger, $\mu_{ci}s$ are spread more evenly among the chosen classes. However, there is no theoretical basis for selecting an optimal m. In this paper, the presented results are obtained, by trial and error, using $m = 2.0$.
2. Calculate the fuzzy cluster centers $\boldsymbol{\Omega}^{(p)}$, using $\mathbf{U}^{(p-1)}$ and (11).
Update $\mathbf{U}^{(p)}$ using $\boldsymbol{\Omega}^{(p)}$ and (12).
3. Compare $\mathbf{U}^{(p)}$ and $\mathbf{U}^{(p-1)}$ with a convenient matrix norm:
If $\Delta\mathbf{U} = \|\mathbf{U}^{(p)} - \mathbf{U}^{(p-1)}\| \le \varepsilon$, then terminate; otherwise set $p = p + 1$ and return to step 2.

B. Supervised Learning Neural Network Implementation

In the search for the minima of the cost function in (7), it is highly efficient to apply a neural network algorithm. The major disadvantage of the neural network however, is that for most users it represents a black box. Nevertheless, extensive research [21] has shown that it is a powerful tool for handling complex problems involving bulky volume set data in high-dimensional feature space. A neural network, combined with fuzzy logic, has also been devised for remote sensing image classification [21]. In DLNN [12], the input–output relationship of the network can be generally expressed as

$$\mathbf{y} = \mathbf{W}\mathbf{x} \tag{13}$$

where the output vector $\mathbf{y}$ contains all the output nodes; $\mathbf{x}$ is a long vector containing all the input and hidden nodes in a network; and matrix $\mathbf{W}$ is formed by concatenating all the weights connected to the output nodes. This modified perceptron structure allows the use of the Kalman filtering algorithm to update the weights during the learning process. This is important, because the implementation of the standard Kalman filter ensures the fast computation of the necessary parameters during the course of learning, namely, the weight updating. Weight updating affects the speed of network learning, which is a key concern when applying the neural network. Weight updating is based on the following criterion

$$\sum_{c=1}^{M} \sum_{i=1}^{N_t} \left\|\mathbf{y}_{ci} - \mathbf{w}_c^{N+1}\mathbf{x}^i\right\| < \varepsilon_{\mathrm{DL}} \tag{14}$$

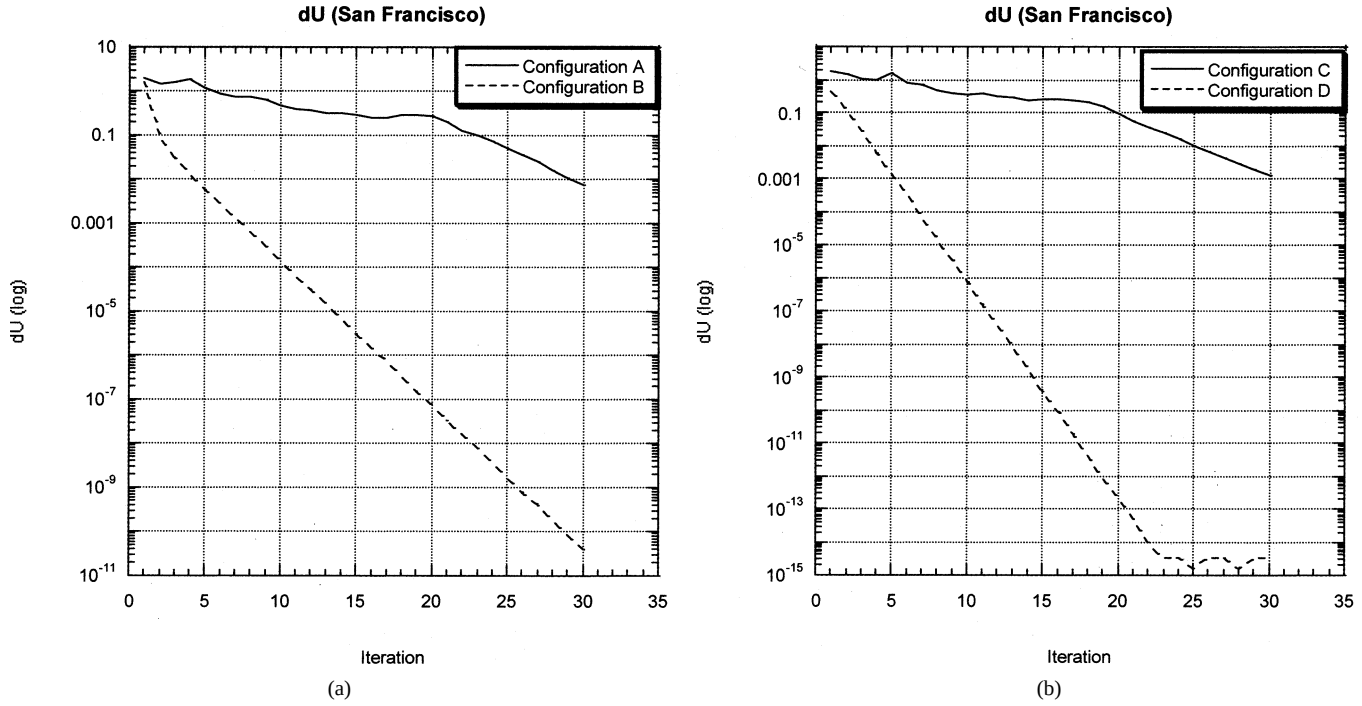


Fig. 5. Plots of matrix norm of $\Delta \mathbf{U} = \|\mathbf{U}^{(p)} - \mathbf{U}^{(p-1)}\|$. (a) *Setup A* and *B*. (b) *Setup C* and *D* (San Francisco Bay Area).

where M is the number of classes; N_t is the total number of input training patterns; and $\mathbf{y}_{ci}$ is the desired output of class c specified by the supervising users. The error bound, $\varepsilon_{\mathrm{DL}}$, can be determined by the users, say, 1%. In a crisp network, we have $\mathbf{y}_{ci} \in \{0, 1\}$, i.e., the input patterns can only belong to one class, while in the fuzzy network patterns can belong to several classes, to a greater or lesser degree. This is what is required to appropriately represent a mixed pixel. To make the output fuzzy rather than crisp, we can simply replace $\mathbf{y}_{ci}$ by a membership value μ_{ci}, which will be described next. In this way, we do not build the fuzzy system into a single neuron, but rather, into a whole network. The detailed implementation of DLNN can be found in [6], [9], and [12].

C. Data Inputs and Outputs

To take full advantage of polarimetric image data, the covariance matrix $\mathbf{C}$ for each image pixel is used as a feature vector. It should be mentioned at this point that in [20], when single-look polarimetric data were used, the classification suffered from serious speckle noise. Therefore, we use a multilook averaged Muller matrix $\mathbf{z}$ with only partial polarimetric information retained for classification [12]. Note that each element is treated with equal significance, because *a priori* class probability is not available. The feature vector can be coded to increase the Hamming distance for better discrimination. For DLNN, because of the use of the Kalman technique, data coding is less beneficial [12]. To this end, there are total of nine input nodes necessary for the adoption of single-frequency polarimetric image data. Polarimetric information, such as intensity ratios, amplitude ratios, and phase differences between different polarizations, are totally and implicitly contained in the polarimetric covariance matrix. Fig. 2 illustrates the data feeding and output configuration of DLNN for the classification of polarimetric SAR data.

D. Training Process

The key to obtaining accurate classification results by a supervised neural classifier is to train the network with representative training patterns within the problem domain. Numerically, the network training is accomplished by adjusting the weights $\mathbf{W}$ between the output nodes $\mathbf{y}$ and the input nodes $\mathbf{x}$. Note that the training datasets should have minimum redundancy to avoid overtraining and to save the computational time.

IV. Experimental Results

A. Test Dataset Description

Two test images, both acquired by the Jet Propulsion Laboratory's Airborne SAR (AIRSAR), were chosen for demonstration. The first one was an L-band (24 cm in wavelength) image of San Francisco, a well-known fully polarimetric SAR test image. The land-cover at this site has been identified as belonging to four classes, i.e., *Ocean*, *Tree*, *Urban*, or *Grass*. Fig. 3(a) and (b) displays the original and filtered images, respectively, of the first test data, in false color. The image is displayed using logarithms of each linear polarization image to enhance its visibility. The second dataset was acquired during the 2000 Pacrim-II campaign over the Au-Ku plantation in western Taiwan. Ground truth work was conducted when the data was taken. From the site survey, a total of six classes was identified: *Sugar cane A*, *Sugar cane B*, *Bare soil*, *Buildings*, *Wetlands*, and *Waterbodies*. *Sugar cane A* and B were sugar canes of two distinctive growth stages. Fig. 4(a) and (b) displays the original and filtered images, respectively, of the second test data.

B. Classification Results and Discussion

For comparison of the effectiveness in classifications, four different configurations, categorized by combinations of feature

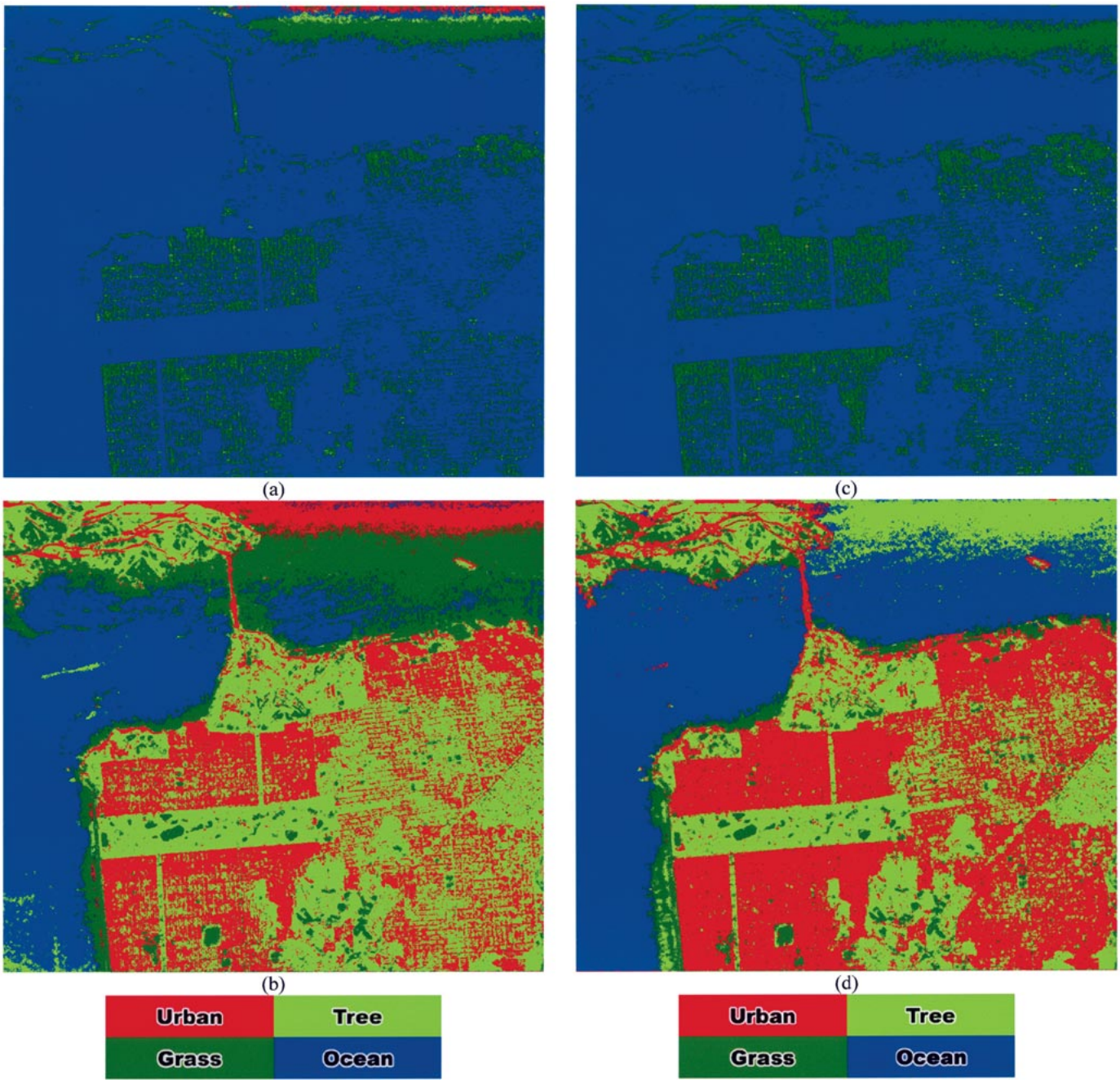


Fig. 6. Classification map of the San Francisco Bay Area. (a) *Setup A*. (b) *Setup B*. (c) *Setup C*. (d) *Setup D*.

vectors and distance measures (see Table I) are devised. The FDL serves as the classifier. *Setups* A and B are formed to examine the applicability of the linear polarizations (HH, HV, and VV) as feature vectors for the purposes of classification in linear and logarithmic scales, respectively. In this case, the complex Wishart distribution distance is not applicable, so the Euclidean distance is used. Note that *setup b* is the same as reported in [12]. Both *setups C* and D use the covariance matrix as the feature vector for classification, but *setup C* uses Euclidean distance for the distance measure, while *setup D* uses the distance based on a complex Wishart distribution. The focus is obviously placed on the competition between *setups C* and D. Note that *Setup D* is the method proposed in this study.

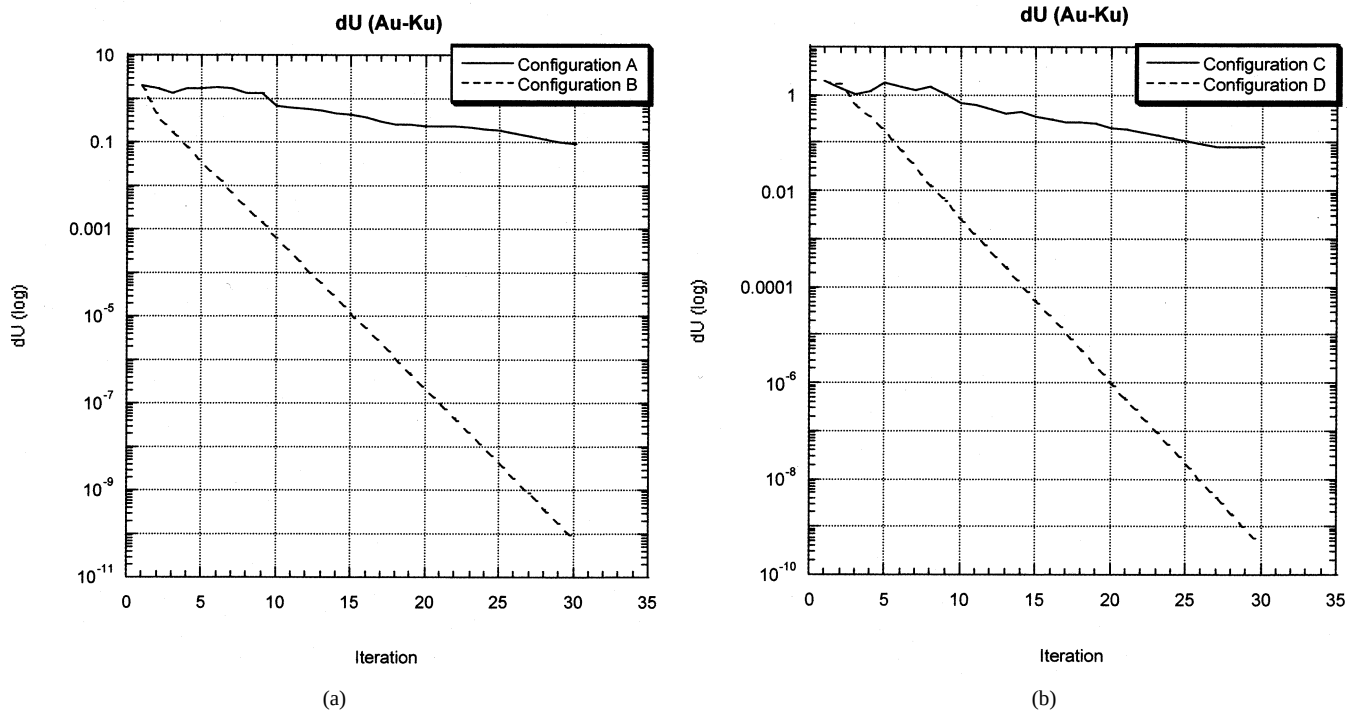
To investigate the performance of these setups, a matrix norm $\Delta\mathbf{U} = \|\mathbf{U}^{(p)} - \mathbf{U}^{(p-1)}\|$, which is the criterion for stopping the fuzzy c-means iteration, is used as a measure of merit. Comparisons are made as to the computation time required to reach convergence. Fig. 5(a) shows the matrix norms for *setups A and B*, while Fig. 5(b) shows them for *setup C* and *D*. Fig. 5(a) clearly shows that the linear polarization data has a faster convergence rate using the logarithmic (decibel) scale (*setup B*) than using the linear scale (*setup A*). In Fig. 5(b), we see that the use of the distance measure based on a complex Wishart distribution (*setup D*) presents a faster convergence rate than the one using the Euclidean distance (*setup C*). A faster convergence rate implies that the Wishart distance measure

TABLE II
CONFUSION MATRIX FOR THE RESULTS OF THE FIRST TEST DATA (SAN FRANCISCO)

		Ocean	Tree	Urban	Grass	Total no. of pixel	Accuracy
Setup A	Ocean	400	0	0	0	400	100.00%
	Tree	400	0	0	0	400	0.00%
	Urban	238	8	1	153	400	0.25%
	Grass	400	0	0	0	400	0.00%
Setup B	Ocean	399	0	1	0	400	99.75%
	Tree	0	397	3	0	400	99.25%
	Urban	0	50	350	0	400	87.50%
	Grass	0	5	0	395	400	98.75%
Setup C	Ocean	400	0	0	0	400	100.00%
	Tree	400	0	0	0	400	0.00%
	Urban	216	7	1	176	400	0.25%
	Grass	400	0	0	0	400	0.00%
Setup D	Ocean	399	0	1	0	400	99.75%
	Tree	0	400	0	0	400	100.00%
	Urban	0	3	393	4	400	98.25%
	Grass	0	3	1	396	400	99.00%

TABLE III
OVERALL TRAINING ACCURACY AND THE KAPPA COEFFICIENT FOR ALL SETUPS. VALUES IN THE PARENTHESIS ARE KAPPA COEFFICIENTS

	Setup A	*Setup B*	*Setup C*	*Setup D*
Test data 1 (San Francisco)	25.06% (0.08%)	96.31% (95.08%)	25.06% (0.08%)	99.25% (99.00%)
Test data 2 (Auku)	40.99% (9.07%)	90.81% (87.57%)	40.99% (9.07%)	91.81% (88.94%)



(a) (b)

Fig. 7. Plots of matrix norm of $\Delta \mathbf{U} = \|\mathbf{U}^{(p)} - \mathbf{U}^{(p-1)}\|$. (a) *Setup A* and *B*. (b) *Setup C* and *D*. The Au-Ku Plantation.

will more precisely measure the distance between features in a nine-dimensional polarimetric covariance matrix space, which makes the fuzzy c-mean clustering converge faster without fluctuations. After comparing the convergence rate, we then inspect the classification results. Fig. 6(a) and (b) displays the classification images for *setups A* and *B*, respectively. From Fig. 6(a), it can be seen that most of the areas are misclassified, indicating that the HH, HV, and VV linear scale data are not suitable as feature vectors without contrast enhancement. *Setup B* shown in Fig. 6(b) has a better classification performance than *setup A*, but only partial polarimetric information is employed leading to greater misclassification. Fig. 6(c) and (d) shows classification images for *setups C* and *D*, respectively. The classification result in Fig. 6(c) is obviously not acceptable, just like the one in Fig. 6(a). This indicates that the Euclidean distance is not suitable for handling the feature vectors from

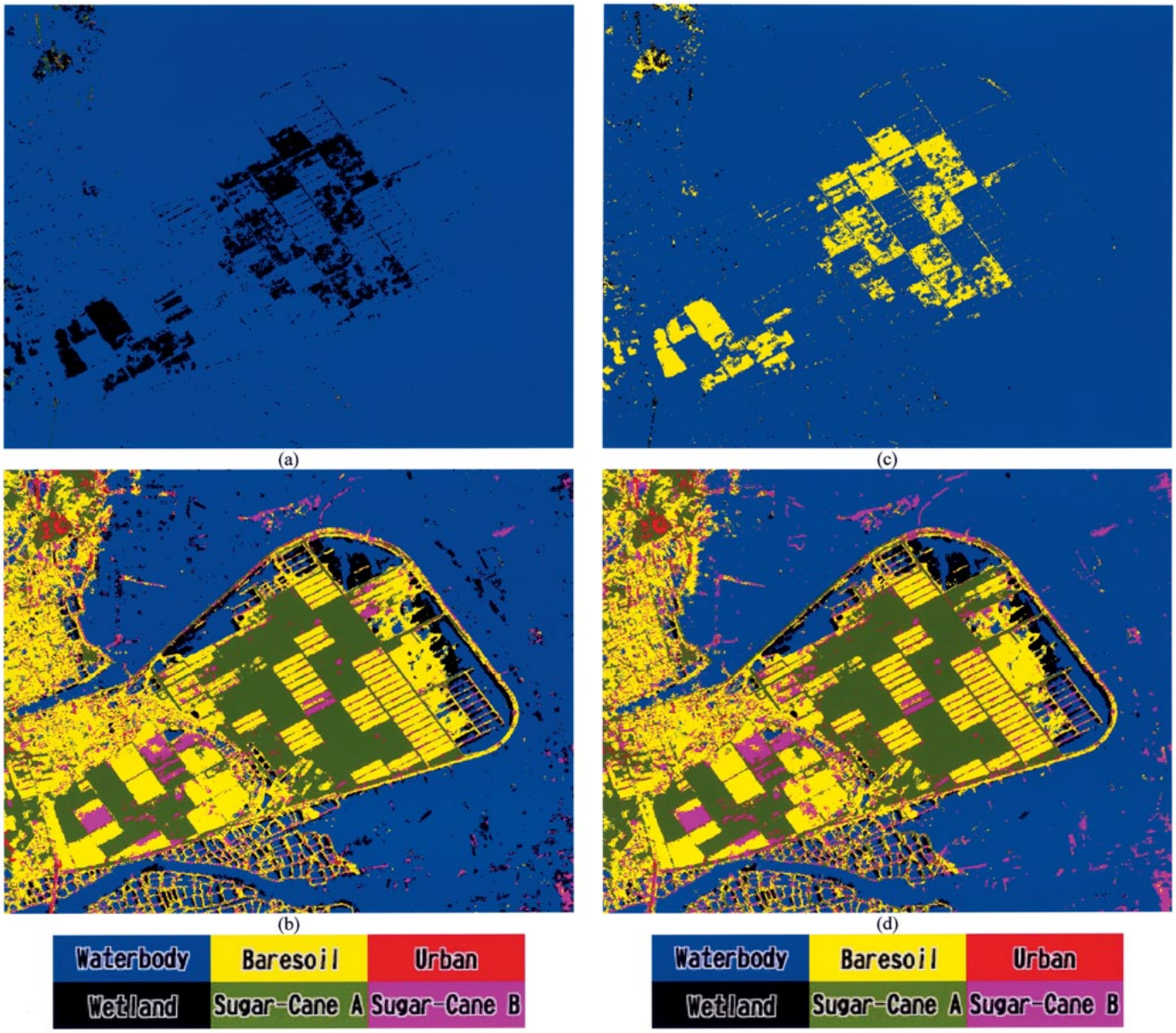


Fig. 8. Classification map of the Au-Ku Plantation. (a) *Setup A*. (b) *Setup B*. (c) *Setup C*. (d) *Setup D*.

fully polarimetric data. The *setup D* classification result, as shown in Fig. 6(c), gives a much better classification accuracy than all other setups (*A*–*C*). Table II shows the confusion table, and Table III shows the overall training accuracy and the kappa coefficients. The confusion table is made by using the check set. We observe that in *setups A* and *C* most of the pixels belonging to the *Tree*, *Urban*, and *Grass* classes are misclassified as *Ocean*, so the training accuracies are obviously unacceptable. It should be mentioned that with *setup C*, we obtain almost equal training accuracy as with *setup A*, even when six more features are employed (off-diagonal terms in the polarimetric covariance matrix). This indicates that additional features do not improve the accuracy unless the appropriate distance measure is applied. The proposed *Setup D* clearly outperforms the other setups, producing better classification accuracy.

The same procedure was also performed on the second dataset, to further validate and confirm the comparisons. Fig. 7 shows the matrix norms for *setups A*, *B*, *C*, and *D*, while the classification results are displayed in Fig. 8. The observations from the first dataset can be equally applied to this dataset. A unique feature of FDL is its capability to regularize the high dynamic range of the input data. For soil surfaces (e.g., the cross-polarized returns usually have value 30 dB below that of the like-polarized returns), the FDL is able to optimally determine the weighting among the components of the given feature vector, to give appropriate input–output mapping, minimizing the numerical error. If necessary, a data encoding scheme may be used to increase the Hamming distance, but this also introduces some information loss. To integrate the amplitude and the phase information representing the different dimensions and dynamic ranges, the neural networks will naturally resolve this in the manner of functional mapping. The generalization capability of the FDL, which enable it to handle complex feature vectors, such as the covariance matrix, is also clear, if the input–outputs can be properly configured.

V. Conclusion

A classification scheme for fully polarimetric SAR image data, which integrates a fuzzy neural network as a classifier, a covariance matrix as a feature vector, and a distance measure based on a complex Wishart distribution, has been demonstrated. Each of these three components plays an essential role in respective aspects to fully explore polarimetric information. The use of FCM during the training phase establishes the capability of handling the mixed pixels for the neural network. This is, the fuzzy neural network enables one to obtain fuzzy partitioning in a supervised way. By employing a complex Wishart distribution distance, the complete polarimetric features can be explored for classification. In practice, polarimetric data for the fuzzy neural network process can be formed as a nine-dimensional feature vector data. We have demonstrated the generalization and regularization capabilities of the proposed method make it very attractive for the processing of fully polarimetric SAR images.

Appendix

To derive the vector of cluster center $\boldsymbol{\omega}_c$ and the membership μ_{ci}, under the complex Wishart distribution, some matrix differential calculus technique needs to be applied. Consider the fuzzy c-means cost function

$$J(\mathbf{U}, \boldsymbol{\Omega}) = \sum_{i=1}^{N_t} \sum_{c=1}^{M} (\mu_{ci})^m \, d^2(\mathbf{z}_i, \boldsymbol{\omega}_c) \tag{A1}$$

and the distance based on the complex Wishart distribution

$$d_w(\mathbf{z}_i, \boldsymbol{\omega}_c) = \ln |\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i). \tag{A2}$$

The optimal fuzzy c-partition of $\mathbf{z}$ is taken as the local minima of the cost function. It is noted that the derivation of the membership function

$$\mu_{ci} = \frac{1}{\sum_{l=1}^{M} \left(\frac{d(\mathbf{z}_i, \boldsymbol{\omega}_c)}{d(\mathbf{z}_i, \omega_l)} \right)^{2/(m-1)}} \tag{A3}$$

is not related to the formation of the distance measure. Hence, the form of the membership function under complex Wishart distribution is the same as the one under a normal distribution (see [22]). Another key parameter is the class center $\boldsymbol{\omega}_c$, derived from the cost function for a given membership function $\mathbf{U}$. Consider the cost function $h(\boldsymbol{\Omega})$

$$\begin{aligned} h(\boldsymbol{\Omega}) &= J(\mathbf{U}, \boldsymbol{\Omega}) \\ &= \sum_{i=1}^{N_t} \sum_{c=1}^{M} (\mu_{ci})^m \, d^2(\mathbf{z}_i, \boldsymbol{\omega}_c) \\ &= \sum_{i=1}^{N_t} \sum_{c=1}^{M} (\mu_{ci})^m \left[\ln |\boldsymbol{\omega}_c| \right. \\ &\quad \left. + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i) \right]^2 \end{aligned} \quad \text{where } \mathbf{U} \text{ is fixed.} \tag{A4}$$

To minimize $h(\boldsymbol{\Omega})$, the partial derivative of $h(\boldsymbol{\Omega})$, with respect to the class center, is set to be zero

$$\frac{\partial h(\boldsymbol{\Omega})}{\partial \boldsymbol{\omega}_c} = \sum_{i=1}^{N_t} \sum_{c=1}^{M} (\mu_{ci})^m \frac{\partial \left[\ln |\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i) \right]^2}{\partial \boldsymbol{\omega}_c} = 0. \tag{A5}$$

For convenience, the function $h(\boldsymbol{\Omega})$ is divided into two parts. One part is dependent only on $\boldsymbol{\omega}_c$

$$f(\boldsymbol{\omega}_c) = \left[\ln |\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i) \right]^2. \tag{A6}$$

Taking the differential of the scalar function f, of matrix $\boldsymbol{\omega}_c$, we obtain

$$\begin{aligned} df(\boldsymbol{\omega}_c) &= 2 \left[\ln |\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i) \right] \, d \left[\ln |\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i \right] \\ &= 2 \left[\ln |\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i) \right] \left\{ \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}) \, d\boldsymbol{\omega}_c \right. \\ &\quad \left. - \mathrm{Tr} \left[\boldsymbol{\omega}_c^{-1} (d\boldsymbol{\omega}_c) \boldsymbol{\omega}_c^{-1} \mathbf{z}_i \right] \right\}. \end{aligned} \tag{A7}$$

Its partial derivatives with respect to class center are

$$\frac{\partial f(\boldsymbol{\omega}_c)}{\partial \boldsymbol{\omega}_c} = 2 \left[\ln |\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i) \right] \left[\boldsymbol{\omega}_c^{-1} - \boldsymbol{\omega}_c^{-1}\boldsymbol{\omega}_c^{-1}\mathbf{z}_i \right]. \tag{A8}$$

Substitute (A8) into (A5), we have

$$\begin{aligned} &\frac{\partial h(\boldsymbol{\Omega})}{\partial \boldsymbol{\omega}_c} = 0 \qquad \forall c \\ &\Rightarrow \sum_{i=1}^{N_t} (\mu_{ci})^m \left[\ln |\boldsymbol{\omega}_c| + \mathrm{Tr}(\boldsymbol{\omega}_c^{-1}\mathbf{z}_i) \right] \\ &\quad \cdot \left[\boldsymbol{\omega}_c^{-1} - \boldsymbol{\omega}_c^{-1}\boldsymbol{\omega}_c^{-1}\mathbf{z}_i \right] = 0 \qquad \forall c \\ &\Rightarrow \sum_{i=1}^{N_t} (\mu_{ci})^m \left[\boldsymbol{\omega}_c^{-1} - \boldsymbol{\omega}_c^{-1}\boldsymbol{\omega}_c^{-1}\mathbf{z}_i \right] d(\mathbf{z}_i, \boldsymbol{\omega}_c) \\ &\quad = 0 \qquad \forall c. \end{aligned} \tag{A9}$$

Following the procedures in [19], if we let $\mathbf{Z}$ have at least $M < N_t$ distinct points, and define $\forall i$ the sets, we have

$$\begin{aligned} I_k &= \{c | 1 \le c \le M;\, d_w(\mathbf{z}_i, \boldsymbol{\omega}_c) = 0\} \\ \tilde{I}_k &= \{1, 2, \ldots, M\} - I_k. \end{aligned}$$

$(\mathbf{U}, \boldsymbol{\Omega}) \in M_{fc}$ may then be globally minimal for J only if

$$I_k = \phi \Rightarrow \mu_{ci} = 1 \Big/ \sum_{l=1}^{M} \left[\frac{d_w(\mathbf{z}_i, \boldsymbol{\omega}_c)}{d_w(\mathbf{z}_i, \boldsymbol{\omega}_l)} \right]^{2/(m-1)} \tag{A10}$$

$$I_k \neq \phi \Rightarrow \mu_{ci} = 0, \qquad \forall c \in \tilde{I}_k \quad \sum_{c \in I_k} \mu_{ci} = 1. \tag{A11}$$

Equation (A11) is the alternate form for the membership of $\mathbf{z}_i$, $\exists c$, so that $d_w(\mathbf{z}_i, \boldsymbol{\omega}_c) = 0$. This is a singularity, and whenever it occurs, z_i must have no membership in any cluster $\boldsymbol{\omega}_c$ in which $d_w(\mathbf{z}_i, \boldsymbol{\omega}_c) > 0$. In practice, care is taken to preclude this eventuality [19]. For $d_w(\mathbf{z}_i, \boldsymbol{\omega}_c) > 0$, (A9) can be written as

$$\sum_{i=1}^{N_t} (\mu_{ci})^m \left[\boldsymbol{\omega}_c^{-1} - \boldsymbol{\omega}_c^{-1}\boldsymbol{\omega}_c^{-1}\mathbf{Z}_i \right] = 0. \tag{A12}$$

Then the cluster center $\boldsymbol{\omega}_c$ is found as follows:

$$\boldsymbol{\omega}_c = \frac{\sum_{i=1}^{N_t} (\mu_{ci})^m \mathbf{Z}_i}{\sum_{i=1}^{N_t} (\mu_{ci})^m}. \tag{A13}$$

As shown in the solution of $\boldsymbol{\omega}_c$, the cluster center vector $\boldsymbol{\omega}_c$, under a complex Wishart distribution, is the same as those for a normal distribution. The results can also be found in [22]. The iterative formulas used for clustering under a complex Wishart distribution are

$$\boldsymbol{\omega}_c = \frac{\sum_{i=1}^{N_t} (\mu_{ci})^m \mathbf{Z}_i}{\sum_{i=1}^{N_t} (\mu_{ci})^m} \qquad \forall\, c$$

$$\mu_{ci} = \frac{1}{\sum_{l=1}^{M} \left(\frac{d(\mathbf{Z}_i, \boldsymbol{\omega}_c)}{d(\mathbf{Z}_i, \boldsymbol{\omega}_l)} \right)^{2/(m-1)}} \qquad \forall\, c,\, i. \tag{A14}$$

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their constructive comments that improve the legibility of this paper.

REFERENCES

[1] J. S. Lee and M. R. Grunes, "Polarimetric SAR speckle filtering and terrain classification—An overview," in *Information Processing for Remote Sensing*, C. H. Chen, Ed, Singapore: World Scientific, 1999, pp. 113–138.

[2] M. R. Azimi-Sadjadi, S. Ghaloum, and R. Zoughi, "Terrain classification in SAR images using principal components analysis and neural network," *IEEE Trans. Geosci. Remote Sensing*, vol. 31, pp. 511–515, Mar. 1993.

[3] Y. Hara, R. G. Atkins, S. H. Yueh, R. T. Shin, and J. A. Kong, "Application of neural networks to radar image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 100–109, Jan. 1994.

[4] J. S. Lee, M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2363–2373, Sept. 1999.

[5] G. M. Foody, M. B. McCulloch, and W. B. Yates, "Crop classification from C-band polarimetric radar data," *Int. J. Remote Sens.*, vol. 15, no. 14, pp. 2871–2885, 1994.

[6] K. S. Chen, W. P. Huang, D. W. Tsay, and F. Amar, "Classification of multifrequency polarimetric SAR image using a dynamic learning neural network," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 814–820, May 1996.

[7] P. D. Heermann and N. Khazene, "Classification of multispectral remote sensing data using a back-propagation neural network," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 81–88, Jan. 1992.

[8] H. Bischof, W. Schneider, and A. H. Pinz, "Multispectral classification of Landsat-image using neural network," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 482–490, May 1992.

[9] K. S. Chen, Y. C. Tzeng, C. F. Chen, and W. L. Kao, "Land-cover classification of multispectral imagery using a dynamic learning neural network," *Photogramm. Eng. Remote Sens.*, vol. 61, no. 4, Apr. 1995.

[10] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy, "Neural network approaches versus statistical methods in classification of multisource remote sensing data," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 540–552, July 1990.

[11] L. Du, J. S. Lee, and S. A. Mango, "Fuzzy classification of earth terrain covers using multi-look polarimetric SAR image data," in *Proc. IGARSS*, vol. 4, 1993, pp. 1602–1604.

[12] Y. C. Tzeng and K. S. Chen, "A fuzzy neural network to SAR image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 301–307, Jan. 1998.

[13] J. S. Lee, K. W. Hoppel, S. M. Mango, and A. R. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 1017–1028, Sept. 1994.

[14] J. J. van Zyl, H. A. Zebker, and C. Elachi, "Imaging radar polarimetric signatures: Theory and observation," *Radio Sci.*, vol. 22, pp. 529–543, 1987.

[15] B. Tayloer, "CYLOPS: The JPL AIRSAR synoptic processor," in *Proc. IGARSS*, Houston, TX, 1992, pp. 652–654.

[16] N. R. Goodman, "Statistical analysis based on a certain complex Gaussian distribution (an introduction)," *Ann. Math. Statist.*, vol. 34, pp. 152–177, 1963.

[17] S. Goze and A. Lopes, "A MMSE speckle filter for full resolution SAR polarimetric data," *J. Electron. Waves Applicat.*, vol. 7, no. 5, pp. 717–737, 1993.

[18] A. Lopes and F. Sery, "Optimal speckle reduction for the product model in multi-look polarimetric SAR imagery and the Wishart distribution," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 632–647, May 1997.

[19] J. C. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*. New York: Plenum, 1987.

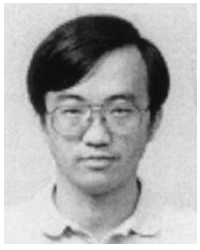
[20] J. S. Lee, M. R. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution," *Int. J. Remote Sens.*, vol. 15, no. 11, pp. 2299–2311, 1994.

[21] F. Wang, "Fuzzy supervised classification of remote sensing images," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 194–201, Mar. 1990.

[22] L. Du and J. S. Lee, "Fuzzy classification of earth terrain covers using complex polarimetric SAR data," *Int. J. Remote Sens.*, vol. 17, no. 4, pp. 809–826, 1996.

Chia-Tang Chen (S'97–M'03) received the B.S. degree in physics from the Chinese Culture University, Taipei, Taiwan, R.O.C., in 1995, and the M.S. and Ph.D. degrees in space science from the National Central University, Chung-Li, Taiwan, R.O.C., in 1997 and 2002, respectively.

He is currently an Assistant Professor with the Department of Information Management, Hsing-Wu College, Taipei, Taiwan, R.O.C. His research interests include signal processing and applications of SAR, image processing, and pattern recognition.



Kun-Shan Chen (S'89–M'90–SM'98) received the B.S. degree from the National Taiwan Institute of Technology, Taipei, Taiwan, R.O.C., in 1985, and the M.S. and Ph.D. degrees from the University of Texas, Arlington, in 1987 and 1990, respectively, all in electrical engineering.

From 1985 to 1990, he was with the Wave Scattering Research Institute, University of Texas. In 1992, he joined the faculty of the Center for Space and Remote Sensing Research, National Central University, Chung-Li, Taiwan, R.O.C., where he is now a Professor and Director. He has joint appointments at the Institute of Space Sciences and Institute of Communication Engineering at the same university. His research activities involve in the areas of microwave remote sensing, image processing and analysis for satellite and aircraft remote sensing data, radio and microwave propagation, and scattering from terrain and ocean with applications to remote sensing and wireless communications. He has authored three book chapters, over 60 refereed journal papers, and over 100 conference papers in the areas of remote sensing and wave scattering and propagation. He has been the Editor-in-Chief of *Journal of Photogrammetry and Remote Sensing* since 2001 and is on the Editorial Board of the *Journal of Electromagnetic Waves and Applications* and *Transactions of the Aeronautical and Astronautical Society of the Republic of China*. He serves as technical consultant at several national research agencies in areas of satellite remote sensing, radar, and radio techniques.

Dr. Chen was the recipient of the 1993 Young Scientist Award from the International Union of Radio Science (URSI) and has received numerous research awards from the National Science Council of Taiwan since 1993. He has been an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. In 2001, he was appointed as chairman of Commission F, Taipei, China of URSI. He is a member of the Electromagnetic Academy. He was a Technical Chairman of PIERS 1999, held in Taipei, Taiwan, R.O.C.

Jong-Sen Lee (S'66–M'69–SM'91–F'97) received the B.S. degree from the National Cheng-Kung University, Tainan, Taiwan, R.O.C., in 1963, and the M.A. and Ph.D. degrees from Harvard University, Cambridge, MA.

He is Head of the Image Science Section, Remote Sensing Division, Naval Research Laboratory, Washinton, DC, where he is the Principal Investigator for several remote sensing programs on polarimetric SAR and interferometric SAR. He has developed several speckle filtering algorithms that have been implemented in many GIS, such as ERDAS, PCI, and ENVI, among others. His research covers a wide spectrum of areas, from control theory, operation research, and radiative transfer to SAR and polarimetric SAR image processing. He has investigated SAR image segmentation, inverse SAR, polarimetric SAR imagery statistics and speckle filtering, SAR polarimetry, and terrain/land-use classification and applications. He was granted a U.S. patent for the invention of a topography measurement technique using polarimetric SAR. His current research interests are in the area of SAR polarimetry, scattering signature modeling, polarimetric SAR interferometry, speckle filtering, and unsupervised classification. He has published more than 50 papers in refereed journals and more than 100 papers in conference proceedings and has given tutorials at IGARSS'97 and IGARSS'98.

Dr. Lee was presented the Best Paper Award and the Best Poster Award at the Third and Fourth European Conference on Synthetic Aperture Radar (EUSAR2000 and EUSAR2002), respectively. He has chaired and organized many sessions in international conferences and is currently an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.