Fully Polarimetric SAR Image Classification Using Different Learning Approaches

Michelle M. Horta
Computer Departament
Federal University of São Carlos, UFSCar
São Carlos, Brazil
michellemh@gmail.com

Nelson D. A. Mascarenhas
Computer Departament
Federal University of São Carlos, UFSCar
São Carlos, Brazil
nelson@dc.ufscar.br

Abstract— This paper compares multilook Polarimetric SAR (PolSAR) image classification using three types of learning: a supervised, an unsupervised and a semi-supervised. The multilook PolSAR pixel values are complex covariance matrices and they are described by mixtures of Wishart distributions. Tests in synthetic and real images showed that the supervised and semi-supervised classifications provided the best results.

Keywords-Multilook Polarimetric SAR Image; Image Classification; Supervised Learning, Unsupervised Learning; Semi-supervised Learning; EM algorithm.

I. INTRODUCTION

The multilook PolSAR imagery has become an important research in Remote Sensing applications [1, 2]. Each pixel is a complex covariance matrix describing the scattering properties of the scene. The drawback of SAR images is the occurrence of multiplicative random speckle noise that degrades the image interpretation and analysis [1].

Focusing on image classification purposes, many statistical techniques have been proposed using different types of learning. However, the Wishart distribution has been used in most applications [1]. In [2], a supervised learning method was proposed through a contextual maximum likelihood algorithm. In [3], unsupervised learning methods were compared using Fuzzy and expectation-maximization clustering. In [4], a stochastic expectation-maximization clustering was proposed.

A recent research [5] applied a semi-supervised learning method to the multilook PolSAR image classification. This semi-supervised learning method uses a determinist annealing clustering technique combined with a multi-layer perceptron classification algorithm.

In multilook PolSAR images applications, the semi-supervised learning is an innovative research topic. Therefore, the main contribution of this paper is a comparison of three types of statistical classifiers using different learning approaches. In all cases, the multilook PoLSAR data are described by a mixture model of Wishart distributions. The supervised learning method is defined by a maximum likelihood classifier. The expectation-maximization clustering proposed in [3] was applied in the unsupervised classification. In addition, this unsupervised method was adapted in order to be used in the semi-supervised classification.

The remainder of this paper is organized as follows: section II defines the Wishart mixture model, section III describes the classification methods, sections IV and V present the experimental results and conclusions, respectively.

II. POLSAR IMAGE DATA

PolSAR imagery is built from polarimetric radar return, which is related to the dielectric properties of the scene. The data are formed by a complex scattering vector $s = (S_{hh}, S_{hv}, S_{vv})$, where (hh, hv, vv) are the polarizations [1]. In a multilook PolSAR image, each pixel is a η-look covariance matrix z given by

$$z = \frac{1}{\eta}\sum_{i=1}^{\eta} s_i s_i^{*T} = \begin{pmatrix} z_{hh} & z_{hhhv} & z_{hhvv} \\ z_{hhhv}^{*} & z_{hv} & z_{hvvv} \\ z_{hhvv}^{*} & z_{hvvv}^{*} & z_{vv} \end{pmatrix}, \qquad (1)$$

where η is the number of looks, $*$ is the conjugate and T denotes transposition. The diagonal elements of the matrix z describe the multilook intensities of the PolSAR data.

In this paper, the pixel datum z obeys a finite mixture model $Z \sim M(\theta,\rho)$, with probability density function defined by $f(z) = \sum_{j=1}^{g} \rho_j f_Z(z,\theta_j)$, where $\sum_{j=1}^{g} \rho_j = 1$ are non-negative proportions, $f_Z(z,\theta_j)$ are the densities, $\theta = (\theta_1,\ldots,\theta_g)$ and $\rho = (\rho_1,\ldots,\rho_{g-1})$ are the parameters and g is the number of distributions [6]. Each jth parameter set (θ_j, ρ_j) of the mixture corresponds to one jth class.

In the mixture of Wishart distributions, the density $f_Z(z,\theta_j)$ characterizes the $W(C_j,\eta)$ distributions and is defined by:

$$f_Z\left(z,\theta_j\right)=\frac{\eta^{\eta m}\left|z\right|^{\eta-m}}{h\left(\eta,m\right)\left|C_j\right|^{\eta}}\exp\left(-\eta Tr\left(C_j^{-1}z\right)\right), \eta\geq m, \quad (2)$$

where $h(\eta,m)=\pi^{m(m-1)/2}\Gamma(\eta)\ldots\Gamma(\eta-m+1)$, $m=3$ is the number of polarizations (hh,hv,vv), Tr and $|\cdot|$ are the trace and the determinant, respectively. The parameter C is a complex covariance matrix as defined in (1) and the expected value of the random variable Z is C [1, 2].

III. Classification Methods

We applied the multilook PolSAR image to three classification approaches: the supervised learning through a maximum likelihood classifier; the unsupervised learning using an expectation-maximization classifier; the semi-supervised learning through an adaptation of the expectation-maximization classifier.

As the image data set obeys a mixture model of Wishart distributions, the log-likelihood function and, consequently, the parameter estimation method are determined by the learning approach [7]. In the supervised learning, the data are labeled, i.e., the classes of the observations are known. Thus, the mixture model can be fitted using the maximum likelihood estimation.

In the unsupervised learning, the classes are not available to fit the mixture model. In the semi-supervised learning, the data set contains labeled and unlabeled data. As a consequence, in both situations, an approximate maximum likelihood algorithm must be used.

A. *Supervised Learning*

This learning approach comprises a training set whereby the classes of the observations are known in advance. Thus, the data set can be represented as $Z^S=\left\{\left(z_1,c_1\right),\ldots,\left(z_N,c_N\right)\right\}$, where z_i is the ith observation and c_i is the corresponding class. As commented previously, each observation z_i is a complex covariance matrix (1).

As the observations are labeled, the maximum likelihood (ML) classification can be used [7,8]. In [2], a contextual maximum likelihood method is proposed to classify multifrequency fully PolSAR images.

Within the mixture model context, the ML classifier can be adapted as follows:

1. Training Step: Define equally probable proportions. Calculate the covariance matrix of the ℓth class by $\hat{C}_\ell=m_1(\tilde{Z})$, where m_1 is the first sample moment and $\tilde{Z}=\left\{z\in Z^S \mid \left(z,c=\ell\right)\right\}$;
2. ML Classifier: Calculate the posterior probabilities for each pixel $z\notin Z^S$ and mixture component by (3). Then, classify the pixel using a maximum a posteriori decision rule.

B. *Unsupervised Learning*

This learning approach consist of only the unlabeled data. Hence, the data set can be represented as $Z^U=\left\{z_1,\ldots,z_N\right\}$, where each ith observation z_i is a complex covariance matrix (1). With such data set, the maximum likelihood estimator of the mixture model is difficult to be computed. This implies that an approximate maximum likelihood algorithm must be used [7].

The expectation-maximization (EM) is an iterative algorithm that formalizes the problem of parameters estimation of a mixture distribution as an incomplete data problem. This idea has often been explored in image clustering applications, as well as in unsupervised image classification [3,6].

The EM clustering using the mixture of Wishart distributions was proposed by [3]. The algorithm can be summarized as follows:

1. Initialization: Define equally probable proportions. Randomly initialize the parameters estimates $\left(\hat{C}_1^{(0)},\ldots,\hat{C}_g^{(0)}\right)$;
2. EM Classifier:

2.1. E-STEP: Update the posterior probabilities for each pixel and mixture component

$$\tau_{ij}^{(k)}=\rho_j^{(k)}f_Z\left(z_i,\theta_j^{(k)}\right)\Big/\sum\nolimits_{\ell=1}^{g}\rho_\ell^{(k)}f_Z\left(z_i,\theta_\ell^{(k)}\right), \quad (3)$$

in which the jth class probability density function is defined by (2);

2.2. M-STEP: Update the parameters estimates by

$$\hat{\rho}_j^{(k+1)}=\frac{\sum_{i=1}^{N}\tau_{ij}^{(k)}}{N}, \quad \hat{C}_j^{(k+1)}=\frac{\sum_{i=1}^{N}\tau_{ij}^{(k)}z_i}{\sum_{i=1}^{N}\tau_{ij}^{(k)}}. \quad (4)$$

The algorithm alternates between these two steps until convergence is achieved. The image can be classified using, for instance, a maximum a posteriori decision rule.

C. *Semi-supervised Learning*

In this learning approach, the data set comprises labeled and unlabeled data, i.e., $Z^{SS}=\left\{Z^{SSU},Z^{SSL}\right\}$ where $Z^{SSU}=\left\{z_1,\ldots,z_m\right\}$ and $Z^{SSL}=\left\{\left(z_{m+1},c_{m+1}\right),\ldots,\left(z_N,c_N\right)\right\}$.

With this data set, the expectation-maximization algorithm can also be used in the semi-supervised learning in order to fit the mixture model [7]. In [9], there is an application of a semi-supervised EM method suitable for image classification.

In this paper, the semi-supervised EM algorithm was applied to multilook PolSAR images. Then, the EM clustering using the mixture of Wishart distributions was adapted as follows:

1. Training Step: Calculate the mixture parameters estimates $\left(\hat{\rho}_j^{(0)}, \hat{C}_j^{(0)}\right)$ using the ML training step and the labeled data Z^{SSL};

2. SSEM Classification:

2.1. E-STEP: Update the posterior probabilities for each unlabeled datum $z_i \in Z^{SSU}$ by (3);

2.2. M-STEP: Update the mixture parameters estimates by (4) and using all the data set Z^{SS}.

IV. RESULTS

We used a synthetic and a real fully polarimetric SAR image to compare the results of three different classification approaches. The real image consists of three classes while the synthetic image comprises five classes.

A. *Real Image*

The real image is an *L*-band image of San Francisco, CA, obtained by the AIRSAR sensor with 4 nominal number of looks. Fig. 1 shows the four pictures used in the analysis: the aerial photo; the sum of the intensities channels (SPAN) and the images including the training and testing areas.

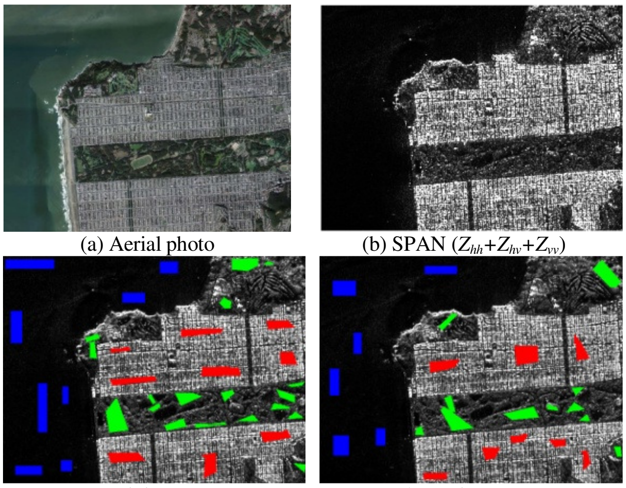
The real image consists of three main classes: sea, vegetation and urban. These classes represent the dark, gray and light areas of the SPAN picture, respectively. The classes have different features in the scene: the sea class describes targets having homogeneous features, the vegetation class describes areas having heterogeneous features and the urban class describes targets having extremely heterogeneous features.

Table I presents the intensity data of the training areas for each class. Although the intensity values of the training data lead to the three separated classes, the vegetation and the urban class have some regions with similar features.

In the Wishart mixture model, the equivalent number of looks was previously estimated ($\hat{\eta} = 3.4$) using data from the sea class. In addition, the number of classes was informed in advance (*g=3*).

Table II shows the kappa coefficient of agreement and the overall accuracy of the classifications. The unsupervised classification results were also analyzed by a supervised manner. The supervised (ML) and semi-supervised (SSEM) classifications were performed using all the training data to estimate the initial parameters, as well as only a percentage of them. The unsupervised (EM) classification was randomly initialized. The SSEM and EM classification results were obtained at iteration 32.

We verified that for this classification problem, the ML method obtained the best results. The EM and SSEM classification results are equivalent. In other words, the kappa coefficients and its variances indicated that the EM and SSEM classification results are not statistically significant. However, the SSEM method provided a faster convergence due to the training data. Fig. 2 depicts the convergence curves of these two classifications.



(a) Aerial photo (b) SPAN ($Z_{hh}+Z_{hv}+Z_{vv}$)

(c) Training areas (d) Testing areas

Figure 1. San Francisco Image (450×600 pixels).

TABLE I. TRAINING DATA OF THE REAL IMAGE

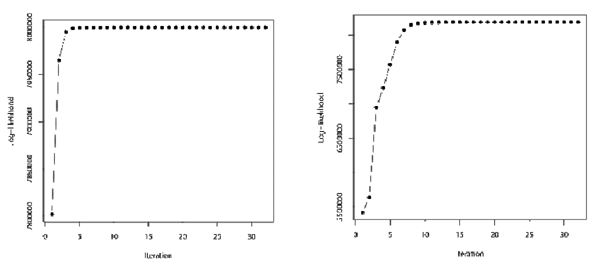
Classes	Intensity Data		
	Z_{hh}	Z_{hv}	Z_{vv}
Sea	0.0044	0.0006	0.0182
Vegetation	0.0517	0.0398	0.0643
Urban	0.2371	0.0531	0.1958

TABLE II. CLASSIFICATION RESULTS OF THE REAL IMAGE

Classification Method	Accuracy	Kappa	Kappa Variance (10^{-5})
ML with 100% TD[a]	83%	0.74	1.8
ML with 25% TD[b]	83%	0.74	1.8
SSEM with 100% TD[a]	77%	0.65	2.2
SSEM with 25% TD[b]	76%	0.65	2.2
EM	76%	0.64	2.2

a. All the training data.

b. Percentage of the training data.



(a) SSEM with 25% TD (b) EM

Figure 2. Convergence curves of the log-likelihood calculated at each iteration.

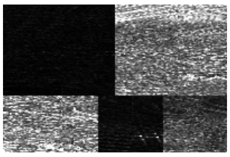
B. *Synthetic Image*

Fig. 3 shows three pictures of the synthetic image: the sum of intensities channels (SPAN) and the images including the training and testing areas. This image consists of five classes. Two classes describe homogeneous areas, which are associated with the blue and cyan regions of the testing areas. The other three classes describe heterogeneous regions. Therefore, there are some classes with close means as shown in Table III.

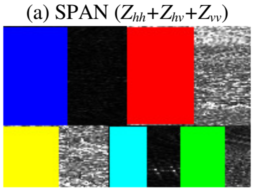
In the Wishart mixture model, the equivalent number of looks was previously estimated ($\hat{n} = 3.4$) using data from the blue class. Similar to the real image, the number of classes was also informed in advance ($g=5$).

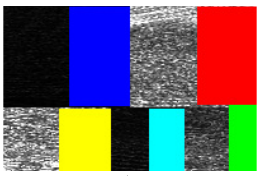
Table IV presents the kappa coefficient of agreement and the overall accuracy of the classifications. In the same way as occurred in the real image situation, the ML and SSEM classifications were performed using all the training data and 25% of them. The EM classification was randomly initialized and the SSEM and EM classification results were obtained at iteration 32.

In this classification problem, the semi-supervised classification performed the best results. The kappa coefficient of the SSEM method with all of training data is higher than the ones obtained in the supervised classifications. The ML classifications and the SSEM classification with 25% of training data are not statistically significant. The unsupervised classification obtained the worst result. Fig. 4 shows the convergence curves of two classifications.



(a) SPAN (Z_{hh}+Z_{hv}+Z_{vv})





(b) Training areas (c) Testing areas

Figure 3. Synthetic Image (100×300 pixels).

TABLE III. TRAINING DATA OF THE SYNTHETIC IMAGE

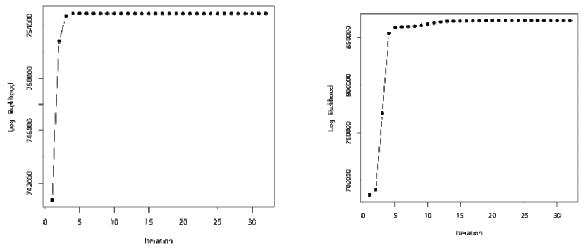
Classes	Intensity Data		
	Z_{hh}	Z_{hv}	Z_{vv}
Blue	0.0053	0.0010	0.0148
Cyan	0.0121	0.0011	0.0128
Red	0.1037	0.0318	0.0897
Yellow	0.1175	0.0133	0.1373
Green	0.0747	0.0141	0.0365

TABLE IV. CLASSIFICATION RESULTS OF THE SYNTHETIC IMAGE

Classification Method	Accuracy	Kappa	Kappa Variance (10^{-5})
ML with 100% TD[a]	82%	0.77	2.0
ML with 25% TD[b]	82%	0.77	2.0
SSEM with 100% TD[a]	85%	0.80	1.7
SSEM with 25% TD[b]	82%	0.76	2.0
EM	78%	0.71	2.2

c. All the training data.

d. Percentage of the training data.



(a) SSEM with 25% TD (b) EM

Figure 4. Convergence curves of the log-likelihood calculated at each iteration.

V. CONCLUSION

This paper analyzed different learning approaches in the context of the multilook PolSAR image classification. The real image tests showed that the supervised classification provided the best results. The unsupervised and semi-supervised accuracies were equivalent. However, the semi-supervised method converged faster than the unsupervised. For the synthetic image experiments, the semi-supervised classification obtained the best results while the unsupervised obtained the worst result. In the synthetic image results, the training data were not enough to fit the distributions of some classes. Therefore, the unlabeled data used by the semi-supervised learning provided more information in the class distributions fitting, making its performance better.

ACKNOWLEDGMENT

This work was supported by FAPESP under grant number 2009/14270-4.

REFERENCES

[1] J. S. Lee, and E. Pottier, Polarimetric radar imaging: from basics to applications. Boca Raton: CRC Press, 2009.

[2] A. C. Frery, A. H. Correia and C. C. Freitas, “Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information,” IEEE Trans. on Geoscience and Remote Sensing, vol. 45, n. 10, pp. 3098–3109, 2007.

[3] P. R. Kersten, J-S. Lee and T. L. Ainsworth, “Unsupervised aperture radar images using fuzzy clustering and EM clustering,” IEEE Trans. on Geoscience and Remote Sensing, vol. 43, n. 3, pp. 519–527, 2005.

[4] M. M. Horta, N. D. A. Mascarenhas, A. C. Frery, and A. L. M. Levada, “Clustering of fully polarimetric SAR data using finite G^0_p mixture model and SEM algorithm,” Proceedings of the 15th International Conference on Systems, Signals and Image Processing, pp. 81-84, 2009.

[5] R. Hänsch, and O. Hellwich, “Semi-supervised learning for classification of polarimetric SAR data,” Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, 2009.

[6] G. MacLachlan, and D. Peel, A Finite Mixture Model. New York: John Wiley & Sons, 2000.

[7] E. Côme, L. Oukhellou, T. Denoeux and P. Aknin, “Learning from partially supervised data using mixture models and belief functions,” Pattern Recognition, vol. 42, pp. 334–348, 2009.

[8] R. O. Duda, P. E. Hart, and D. G. Stock Pattern Classification, 2nd ed., New York: John Wiley & Sons, 2001.

[9] A. Baraldi, L. Bruzzone and P. Blonda, “A multiscale expectation-maximization semi-supervised classifier suitable for badly posed image classification,” IEEE Trans. on Image Processing, vol. 15, n. 8, pp. 2208–2224, 2006.