Unsupervised Classification of Remote Sensing Data using Graph Cut-Based Initialization*

Mayank Tyagi*, Ankit K Mehra*, Subhasis Chaudhuri*
and Lorenzo Bruzzone** {(mayanktyagi, ankitmehra, sc)@iitb.ac.in,
lorenzo.bruzzone@ing.unitn.it}

* IIT-Bombay, India. ** University of Trento, Italy.

Abstract. In this paper we propose a multistage unsupervised classifier which uses graph-cut to produce initial segments which are made up of pixels with similar spectral properties, subsequently labelled by a fuzzy c-means clustering algorithm into a known number of classes. These initial segmentation results are used as a seed to the expectation maximization (EM) algorithm. Final classification map is produced by using the maximum likelihood (ML) classifier, performance of which is quite good as compared to other unsupervised classification techniques.

1 Introduction

In literature, two main approaches to the classification problem have been proposed: the supervised approach and the unsupervised approach [1]. The supervised classification methods require the availability of a suitable ground truth (training data) for the classification. On the other hand, the pattern classification using unsupervised techniques can be performed by only considering the information contained in the image itself. Normally, all supervised methods offer a higher accuracy compared to unsupervised classification methods. Unfortunately, in many applications, we do not have ground truth information. In fact, gathering a sufficient number of training samples for each specific image considered, by either photo-interpretation or through the collection of ground truth information, is very expensive in terms of time and economy. Also we have to regularly update this ground truth information since atmospheric conditions and hence digital signatures of different classes change with time. Therefore, in many cases, it is not possible to obtain training data for each image to be classified. So we need good unsupervised methods to perform classification in such cases.

In this paper, we propose a novel multistage unsupervised classifier, based on the maximum likelihood (ML) criterion. This classifier exploits both the region-wise and pixel-wise spectral information included in the image to give reliable classification results. The classical unsupervised classifiers like fuzzy c-means [2] [3], and classifiers using artificial neural network techniques [4] [5] are pixel-based classifiers and do not consider any parametric modeling of the population distribution of a class. We present here a hierarchical approach, in

* Funding under Indo-Trento ITPAR project is acknowledged.

which classification is performed on segments generated by graph cut based segmentation, instead of performing on pixel-level. Our approach is based on graph theory based image segmentation and fuzzy classification technique. These initial classification results are used as a seed to an expectation maximization (EM) algorithm. Experimental analysis is carried out on a data set related to the Island of Sardinia (Italy). The results obtained by the proposed unsupervised classifier confirms the effectiveness of the proposed approach.

2 Problem Formulation

Let $\mathbf{X} = \{\mathbf{x}_{1,1}, \mathbf{x}_{1,2}, \ldots, \mathbf{x}_{R,S}\}$ denote a multi spectral image composed of $R \times S$ pixels. Let $\mathbf{x}_{r,s}$ be the $(1 \times j)$ feature vector associated with the pixel at position (r, s) of the image $\mathbf{X}$ (where j is the number of spectral bands present). Let us assume that the set $\Omega = \{\omega_1, \omega_2, \ldots, \omega_M\}$ of M land-cover classes characterizes the considered geographical area in image $\mathbf{X}$. Moreover, it is assumed that the total number of land-cover classes (M) in the set are known *a priori*. Let $l_{r,s}$ be the class label of pixel $\mathbf{x}_{r,s}$.

In the context of the Bayes decision theory [7], the decision rule adopted by the ML classifier for classification is expressed as follows:

$$l_{r,s} \in \omega_m, \text{ if } \omega_m = \arg \max_{\omega_i \in \Omega} \{\hat{P}(\omega_i)\hat{p}(\mathbf{x}_{r,s}|\omega_\mathbf{i})\} \tag{1}$$

where $\hat{P}(\omega_i)$ and $\hat{p}(\mathbf{X}|\omega_\mathbf{i})$ are the estimates of the *a priori* probability and the conditional density function of the class ω_i in the image $\mathbf{X}$, respectively. The training phase of an ML classifier consists in the estimations of the *a priori* probability $P(\omega_i)$ and the conditional density function $p(\mathbf{X}|\omega_\mathbf{i})$ for each class $\omega_i \in \Omega$. In all practical classification approaches, these estimates are obtained by using classical supervised methods that exploit the information included in the appropriate training set for the image. However in our case no training data is available and the probabilities have to be estimated from the data itself.

3 Proposed Unsupervised Classifier

Consider the case in which all classes included in Ω, the land-cover classes set, can be described by Gaussian distributions. The density function associated with each class ω_i can be completely described by the mean vector μ_i, and the covariance matrix Σ_i. Therefore, the parameters to be estimated are:

$$\theta = [\mu_1, \Sigma_1, P(\omega_1), \ldots, \mu_M, \Sigma_M, P(\omega_M)] \tag{2}$$

The estimation of the probabilities becomes a mixture density estimation problem. The expectation maximization (EM) algorithm is one of the most powerful solutions to such type of problems [6]. It can be proved that the required equations for the estimation of means and the covariance matrices for different classes in Ω are the following [7]:

$$P^{k+1}(\omega_i) = \frac{\sum_{\mathbf{x}_{r,s}\in\mathbf{X}} \frac{P^k(\omega_i)p^k(\mathbf{x}_{r,s}|\omega_i)}{p^k(\mathbf{x}_{r,s})}}{R \times S} \tag{3}$$

$$[\mu_i]^{k+1} = \frac{\sum_{\mathbf{x}_{r,s}\in\mathbf{X}} \frac{P^k(\omega_i)p^k(\mathbf{x}_{r,s}|\omega_i)}{p^k(\mathbf{x}_{r,s})}\mathbf{x}_{r,s}}{\sum_{\mathbf{x}_{r,s}\in\mathbf{X}} \frac{P^k(\omega_i)p^k(\mathbf{x}_{r,s}|\omega_i)}{p^k(\mathbf{x}_{r,s})}} \tag{4}$$

$$[\Sigma_i]^{k+1} = \frac{\sum_{\mathbf{x}_{r,s}\in\mathbf{X}} \frac{P^k(\omega_i)p^k(\mathbf{x}_{r,s}|\omega_i)}{p^k(\mathbf{x}_{r,s})}\left(\tilde{\mathbf{x}}_{r,s,i}.\tilde{\mathbf{x}}_{r,s,i}^T\right)}{\sum_{\mathbf{x}_{r,s}\in\mathbf{X}} \frac{P^k(\omega_i)p^k(\mathbf{x}_{r,s}|\omega_i)}{p^k(\mathbf{x}_{r,s})}} \tag{5}$$

where

$$p(\mathbf{x}_{r,s}) \ = \sum_{i=1}^{M} P(\omega_i)p(\mathbf{x}_{r,s}|\omega_i) \text{ and } \tilde{\mathbf{x}}_{r,s,i} = \mathbf{x}_{r,s} - [\mu_i]^k \tag{6}$$

k is the iteration index and $R \times S$ is the total number of the pixels in image $\mathbf{X}$.

The EM algorithm is not guaranteed to converge to the global maximum. This is due to the fact that the maximum, whether global or local, to which the algorithm converges depends on the initial seed given for the iterations, i.e. $P^0(\omega_i), \mu^0(\omega_i)$ *and* $\Sigma^0(\omega_i)$. The proposed classifier obtains the seeds for the iterations through an unsupervised method, from the image $\mathbf{X}$ itself, using the graph cut based image segmentation followed by the fuzzy c-means algorithm.

As the first step, the proposed classifier applies the graph-cut based segmentation technique [8] on the image $\mathbf{X}$ to obtain an over segmentation $\mathbf{G}$ from the image $\mathbf{X}$. This segmentation $\mathbf{G}$ consists of components $\mathbf{C_i}$ such that each $\mathbf{C_i}$ is a connected subgraph of the original graph. It is called an over segmentation because even if two components $\mathbf{C_i}$ and $\mathbf{C_j}$ belong to the same class but are not connected spatially, these components will be considered as two different segments. Therefore, the number of components present will be typically much greater than the number of classes present in the image. This segmentation is done on the basis of spectral properties of the image pixels only. The algorithm by Felzenszwalb and Huttenlocher [8], being fast and accurate, is used for this step. Now, as the number of segments into which the image is divided is greater than the number of classes present, the segments are merged together on the basis of similarity between the means and variances of the two segments being considered. That is, a segment $\mathbf{C}$ belonging to $\mathbf{G}'$, segmentation obtained after merging, is obtained from $\mathbf{G}$ by merging two segments $\mathbf{C}_i \in G$ and $\mathbf{C}_j \in G$, if the differences between the means and variances of two segments is less than some thresholds, which is empirically determined.

At the end of the above step, if the final segmentation $\mathbf{G}' = \{C_1, C_2, \ldots, C_N\}$ has N segments where $N > M$ then the total number of segments N has to be reduced to M segments so that it can be used as seeds for the EM algorithm iterations. For this step fuzzy c-means algorithm is used [2]. That is, say, if $\{\mu_1, \mu_2, \ldots, \mu_N\}$ denotes the mean vector of segments in $\mathbf{G}'$, the fuzzy c-means algorithm is applied on these segments to cluster them and obtain M number of new segments with new means $\{\mu'_1, \mu'_2, \ldots, \mu'_M\}$. To take into account the different sizes of the segments while clustering, each pixel value of a new cluster is multiplied by a weight factor, which is directly proportional to the size of its original cluster, while calculating the new mean for the resulting segment.

After obtaining **M** number of segments and their respective means, the covariance matrices and the membership probabilities for each segment are calculated from the clustered pixels. Now as these segments can be considered as the land cover classes for the area, the corresponding means and covariances can be taken as initial estimates for class conditional density function $p(\mathbf{X}|\omega_\mathbf{i})$ parameters, and the membership probability as the *a priori* class probability $P(\omega_i)$ for all **M** land-cover classes. Thus the initial estimate of θ in equation 2 is obtained. These estimates are then given as seed, $P^0(\omega_i)$ and $p^0(\mathbf{X}|\omega_\mathbf{i})$ to the EM algorithm iterations, which calculates the new parameters for the ML classifier, using the spectral information included in image.

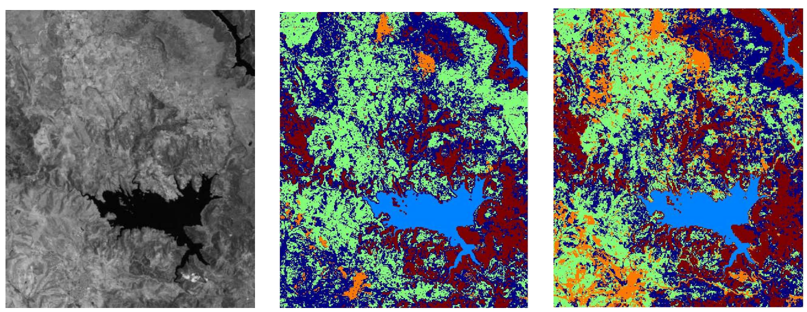
4 Experimental Results

Experiments were carried out on a data set made up of a multi-spectral image acquired by the Thematic Mapper (TM) multi spectral sensor of the Landsat 5 satellite. The selected test site was a section (493×412 p$ixels$) of a scene including Lake Mulargias on the Island of Sardinia, Italy acquired in July 1996. Fig. 1(a) shows band 5 of the image. In particular, five land cover classes (i.e. pasture, forest, urban area, water, vineyard) were identified (by domain expert) that characterized the test site at the above date.

The results obtained using proposed classifier, when applied to the image, are shown in Table I. The table lists the total number of test samples for each class, as well as number of pixels classified accurately by the classifier. As shown in Table I, the total accuracy obtained by the proposed classifier on the data set is about 80.21%. As a comparison,the fuzzy c-means algorithm on the image gave an accuracy of only 70.12%. Although the proposed classifier outperforms the unsupervised fuzzy c-means classifier by about 10%, the overall accuracy obtained was not too high.

In order to understand this, when we studied the distributions (in the feature space) of subset of pixels corresponding to all the classes, we observed that two of the land-cover classes (pasture and vineyard) had very similar spectral characteristics and the two corresponding distributions are getting overlapped. Due to this, pixels belonging to one class are getting inaccurately classified to other class, yielding a low classification accuracy. Also, the vineyards are very small regions in the image and some of these regions were not segmented properly by the graph-cut method, further aggravating the cause.

In order to overcome this problem, pixels belonging to these two classes (pasture and vineyard) were taken to be belonging to one class only. Therefore, the problem reduces to classifying four land-cover classes only. The application of the proposed classifier within such a framework gave an accuracy of 86.33% (as shown in Table II). As a comparison, the corresponding fuzzy c-means classifier gave an overall accuracy of 78.69%, which is notably much lower to that of the proposed classifier. It may be noted that test data set is available only at 1844 pixels, whereas we have classified the entire 493×412 pixels. Since the domain knowledge is not available at all points, the accuracy cannot be computed for the entire land cover. However, a quick look at the classification results shown in Fig. 1(b) suggest that the actual accuracy of the proposed classifier is much higher since it appears to segment a particular class at a given region quite well. In Fig. 1(c), we reproduced the results of supervised classification as developed in [7] using a proper training data, whose accuracy is 95.64%. Compare this to the classification results using the proposed technique given in Fig. 1(b). Given the fact that no training data was provided, the results are still very good. It may be noted that an unsupervised clustering techniques can do segmentation but not the labeling of segments into actual class labels such as waterbody or pasture. The class labels were assigned manually.



(a) (b) (c)

Fig. 1. (a) Band 5 of a Landsat image. Results of (b) proposed classifier and (c) supervised ML classifier [6].

5 Conclusions

A novel unsupervised classifier, based on estimating the unknown classifier parameters by using graph cut based segmentation and fuzzy c-means algorithm, has been developed. The EM algorithm is used to learn the class distribution. The proposed method is different from other unsupervised classifiers in the sense that it exploits both the regional as well as pixel wise spectral information contained in the image, giving quite a high accuracy. Our method gave better results than fuzzy c-means and other unsupervised method but the accuracy is still lower than the supervised methods (normally above 90% compared to our 80%). But

Table 1. Performance comparison of the proposed method for five classes.

Land-Cover Class	Number of test samples	Proposed Unsupervised Classifier(%)	Fuzzy c-means Classifier (%)	Supervised ML Classifier (%)
Pasture	149	46.98	6.71	96.39
Forest	304	63.82	56.56	86.51
Urban Area	408	99.75	75.00	98.78
Water	804	100.00	100.00	100
Vineyard	179	2.23	0.56	82.13
Overall	1844	80.21	70.12	95.64

Table 2. Performance comparison of the proposed method for four classes.

Land-Cover Class	Number of test samples	Proposed Unsupervised Classifier(%)	Fuzzy c-means Classifier (%)	Supervised ML Classifier (%)
Pasture+Vineyard	328	59.45	66.46	94.27
Forest	304	65.13	77.63	86.51
Urban Area	408	96.81	47.30	98.28
Water	804	100	100	100
Overall	1844	86.33	78.69	95.99

these initial results can be very helpful to a domain expert who has to assign classes to training samples. If the domain expert has good initial estimates then he/she will have the convenience of assigning classes mostly at a region level rather than at a pixel level.

References

1. R.O. Duda, P.E. Hart, D.G. Stork "Pattern Classification", *John Wiley & Sons, Inc.*, (2001).
2. R. Ehrlich, J.C. Bezedek, W. Full, "FCM: The Fuzzy C-Means Clustering Algorithm", *Comp. Geoscience, vol.10*, (1984) 191-203.
3. I. Gath, A. B. Geva, "Unsupervised Optimal Fuzzy Clustering", *IEEE PAMI, vol. 11*, (1989) 773-780.
4. S. Haykins, "Neural Networks", *Pearson Education*, (1999).
5. D. Kilpatrick, R. Williams, "Unsupervised Classification of Antarctic Satellite Imagery Using Kohonen's Self-Organising Feature Map", *Proceedings, IEEE International Conference on Neural Networks, vol. 1*, (1995) 32-36.
6. T.K. Moon, "The Expectation-Maximization Algorithm", *IEEE Signal Processing Magazine*, (November 1996) 47-60.
7. L. Bruzzone, D.F. Prieto, "Unsupervised Retraining of a Maximum Likelihood Classifier for the Analysis of Multitemporal Remote Sensing Images", *IEEE Transactions On Geoscience And Remote Sensing, Vol. 39, No. 2*, (February 2001) 456-460.
8. P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient Graph-Based Image Segmentation", *International Journal of Computer Vision, Vol. 59, Number 2* , (September 2004)