SWARM INTELLIGENCE FOR UNSUPERVISED CLASSIFICATION OF HYPERSPECTRAL IMAGES

Andrea Paoli, Farid Melgani, Edoardo Pasolli

Dept. of Information Engineering and Computer Science, Univ. of Trento,
Via Sommarive, 14, I-38100 Trento, Italy
E-mail: melgani@disi.unitn.it

PREFERRED TOPIC: T3 – CLASSIFICATION AND DATA MINING TECHNIQUES

PRESENTATION: ORAL

ABSTRACT

From a methodological viewpoint, a classification process consists of associating a pattern (sample) to a class label opportunely chosen from a predefined set of class labels. In the literature, two main approaches to the classification problem have been proposed: the supervised and the unsupervised approaches. Supervised techniques require the availability of a training set for learning the classifier. Unsupervised methods, known also as clustering methods, perform classification just by exploiting information conveyed by the data, without requiring any training sample set. The supervised methods offer a higher classification accuracy compared to the unsupervised ones, but in some applications it is necessary to resort to unsupervised techniques because training information is not available.

In this work, we focus the attention on hyperspectral image clustering. Compared with conventional multispectral data, hyperspectral data are characterized by a higher spectral resolution giving thus the opportunity to push further the information extraction capability. However, hyperspectral imagery involves a greater quantity of data to memorize and to process. Moreover, given a specific classification problem, hyperspectral data often exhibit redundant information, calling thus for opportune band (feature) selection algorithms. While feature selection has been widely studied in the supervised classification context, little has been done in the image clustering context due to the lack of training samples. Another intrinsic problem in image clustering in general and in hyperspectral image clustering in particular is how to set a priori the number of data classes because of the absence of prior information.

In the literature, there are still few published works dealing with the clustering of remote sensing hyperspectral images. In [1], the hybrid supervised-unsupervised approach to image classification was improved by introducing the concept of cluster-space classification for hyperspectral data. Cluster-space representation is used for associating spectral clusters with corresponding information classes automatically, thus overcoming the manual assignment of clusters and classes carried out in the hybrid approach. In [2], a method of hyperspectral band reduction based on rough sets and fuzzy C-means clustering was proposed. It consists of two steps. First, the fuzzy C-means clustering algorithm is used to classify the original bands into equivalent band groups. Then, data dimensionality is reduced by selecting only the band with maximum grade of fuzzy membership from each of the groups. In [3], the authors presented a two-stage hierarchical clustering technique for classifying hyperspectral data. First, a "local" segmentator performs region-growing segmentation by merging spatially adjacent clusters. Then, a "global" segmentator clusters the segments resulting from the previous stage using an agglomerative hierarchical clustering scheme based on a context-free similarity measure. In [4], a two-step unsupervised artificial immune classifier for multi/hyperspectral images was presented.

In this paper, we propose a novel clustering system for hyperspectral images capable, without any a priori knowledge about the investigated area, of solving simultaneously three different problems: 1)

cluster parameter estimation; 2) selection of the most discriminative features; 3) class number estimation. Given the hyperspectral nature of data, we suppose that samples are normally distributed. In this way, each image pixel takes origin from a multivariate Gaussian distribution described by a mean vector and a covariance matrix assumed to be diagonal in order to simplify the clustering problem. Accordingly, the image is statistically described by a mixture of Gaussian distributions. The cluster parameter estimation issue becomes the one of determining the mixture parameter values that best approximate the image distribution. The feature selection has the purpose to reduce data redundancy and to remove the bands that are characterized by a strong noise component. Both cluster parameter estimation and feature selection issues are simultaneously faced by a multiobjective approach based on particle swarm optimization (PSO) [5]. In an optimization problem formulated within a PSO framework, it is necessary to define the ingredients of the PSO algorithm, namely the particle position and the fitness functions. In our case, the position of each particle is simply a vector that concatenates all the variables to estimate, i.e., the cluster parameters and the features to select. Concerning the fitness functions, we propose to optimize jointly two different criteria: the likelihood function to estimate the cluster parameters and the Bhattacharyya distance to select the most discriminative features. The last issue, i.e., the class number estimation, is handled by repeating the PSO process in a predefined range of values of class number and thus optimizing the minimum description length (MDL) criterion [6]. At the end of the learning process, a classification map is generated by applying the maximum a posteriori (MAP) decision criterion for each image pixel.

To illustrate the performance of the proposed system, we conducted an experimental study based on both simulated and real (AVIRIS and ROSIS) hyperspectral images. Simulated images were generated by varying the number of classes and their prior probabilities, the number of total features and of noisy features, and the noise intensity. The results were compared with those yielded by the k-means algorithm applied before and after a feature reduction step based on principal component analysis. In general, the obtained experimental results show that interesting performances can be achieved though the processing context is completely unsupervised.

Keywords: Feature selection, hyperspectral images, image clustering, k-means algorithm, multiobjective optimization, particle swarm optimization (PSO).

REFERENCES

[1] X. Jia and J. A. Richards, "Cluster-space representation for hyperspectral data classification," *IEEE Trans. Geosci. and Remote Sens.*, vol. 40, no. 3, pp. 593–598, 2002.

[2] H. Shi, Y. Shen, and Z. Liu, "Hyperspectral bands reduction based on rough sets and fuzzy C-means clustering," in *Proc. 20th IEEE Instrumentation and Measurement Technology Conference*, Vail, USA, May 20–22, 2003, vol. 2, pp. 1053–1056.

[3] S. Lee and M. M. Crawford, "Hierarchical clustering approach for unsupervised image classification of hyperspectral data," in *Proc. IEEE Geosci. and Remote Sens. Symp.*, Anchorage, Alaska, Sep. 20–24, 2004, vol. 2, pp. 941–944.

[4] Y. Zhong, L. Zhang, B. Huang, and P. Li, "An unsupervised artificial immune classifier for multi/hyperspectral remote sensing imagery," *IEEE Trans. Geosci. and Remote Sens.*, vol. 44, no. 2, pp. 420–431, 2006.

[5] Y. Bazi and F. Melgani, "Semisupervised PSO-SVM regression for biophysical parameter estimation," *IEEE Trans. Geosci. and Remote Sens.*, vol. 45, no. 6, pp. 1887–1895, 2007.

[6] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, Singapore: World Scientific, 1989.