

Hierarchical Clustering of Self-Organizing Maps for Cloud Classification

Christophe Ambroise^a Fouad Badran^a Sylvie Thiria^b
Geniève Sèze^c

^a *Université Pierre et Marie Curie, 75005 Paris, France*

^b *Conservatoire National des Arts et Métiers, 75003 Paris, France*

^c *École Polytechnique, 91128 Palaiseau Cedex, France*

Abstract

This paper presents a new method for segmenting multispectral satellite images. The proposed method is unsupervised and consists of two steps. During the first step the pixels of a learning set are summarized by a set of codebook vectors using a Probabilistic Self-Organizing Map (PSOM, [9]) In a second step the codebook vectors of the map are clustered using Agglomerative Hierarchical Clustering (AHC, [7]). Each pixel takes the label of its nearest codebook vector. A practical application to Meteosat images illustrates the relevance of our approach.

Key words: Kohonen maps, image segmentation, hierarchical clustering, cloud classification

1 Introduction

Classification of remote sensed imagery is usually supervised. One of the problem with cloud classification is the difficulty of getting a great number of labeled samples which represent the diversity of all possible situations. As supervised learning uses only labeled samples for determining the classifier parameters and for assessing the classifier performance, it can lead to biased results: some classes may be not represented or under sampled.

In this paper a fully unsupervised approach is proposed. The method is based on Probabilistic Self-organizing Maps [9, 1] and works in two steps. The first steps approximates the distribution of the pixels to be classified using a topological map. Hierarchical clustering is then used for labeling the neurons of map.

Section 3 describes the proposed methodology. A real numerical experiment is proposed in section 4. The last section discusses the limits and possible improvements of our classifier.

2 Combining Probabilistic SOM with Hierarchical Clustering

Hierarchical Clustering can be used for classifying several thousand of objects but is unable to deal with huge data set such as meteosat images. Using probabilistic topological algorithms in a first stage allow to summarize the initial data set into a smaller set of codebook vectors which can be clustered using a hierarchical clustering algorithm [6]. The topological algorithm typically provides a partition of thousand clusters, which is used as starting point for hierarchical clustering. Let us detail the two steps of this procedure.

2.1 Probabilistic SOM

Vector quantization algorithms are classically used to produce an approximation to the distribution of a data set. These algorithms find a set of codebook vectors to code compactly the whole vectors of the dataset. Compare to classical vector quantization method, topological maps introduces a additional relation between the codebook vectors. This “topological” relationship is a constraint which has been proved to produce robust results [5].

The most well-known topological algorithm is the Kohonen Self-Organizing Maps Algorithm. It assumes implicitly that the distribution to approximate is a Gaussian mixture where each component is defined by a mean vector (the codebook vector) and a covariance matrix of the form $\Sigma_k = \lambda.I$. All the components are supposed to have the same covariance matrix. Generalisation of this algorithm have recently arised in the framework of this probabilistic interpretation of Kohonen algorithm [2, 9]. For example the authors [9] have considered that the distribution of the vectors $\mathbf{z} \in \mathbb{R}^d$ can be expressed as

$$p(\mathbf{z}) = \sum_{\mathbf{c}} p(\mathbf{c})p_c(\mathbf{z}),$$

where $p_c(\mathbf{z})$ represent a Gaussian Mixture related to the codebook vector \mathbf{c} . The *prior* probability of this mixture are determined from topological distances δ between the codebook vectors on the map :

$$p_c(\mathbf{z}) = \frac{1}{T_c} \sum_r K(\delta(\mathbf{r}, \mathbf{c})) f_r(\mathbf{z}),$$

where $K()$ is a positive kernel function which determines how the codebook vectors are related to each other, $T_c = \sum_r K(\delta(\mathbf{r}, \mathbf{c}))$ is a normalization constant and $f_r(\mathbf{z})$ a normal density. This kind of algorithm aims to find the model parameter which maximizes a given criterion such as the Likelihood.

2.2 Agglomerative Hierarchical Clustering

Agglomerative Hierarchical Clustering is an iterative procedure which is based on the following principle:

- (1) Find the two closest clusters according some dissimilarity.
- (2) Agglomerate them to form a new cluster.

The first iteration considers the partition induced by the codebook vectors of the map. The next iterations require to define a dissimilarity between the merged clusters and all other clusters. There are many ways to do so.

The most widely considered rule for merging clusters may be the method proposed by Ward [7]. This method measures the cluster dissimilarity using a ponderated squared distance between the two cluster centers μ_{ell} and μ_k :

$$\frac{n_k \cdot n_\ell}{n_k + n_\ell} \|\mu_k - \mu_\ell\|^2,$$

where n_k and n_ℓ represents the cardinalities of clusters k and ℓ . This distance has been used in this paper.

The labeled map can then be used for classifying new set of objects issued from similar measurment

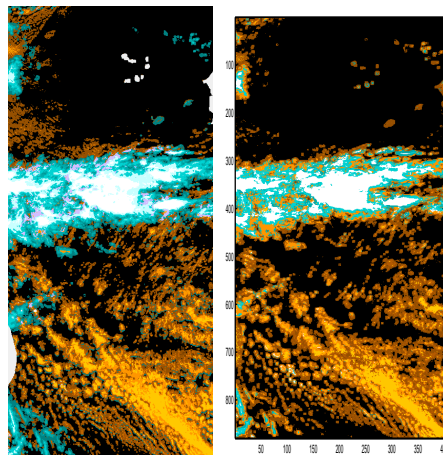
3 Cloud Classification Experiment

The data set we consider here, is extracted from a larger data set built in the frame of the POLDER experiment. This larger data set provided by METEO FRANCE consists in METEOSAT full field view visible and infrared radiance field, at 5km resolution (full IR resolution, VIS sampled one pixel other two) and one half hour time period between 6h GMT and 15h GMT for the 30 October 10 November 1996 period. This data set has been calibrated and the visible radiance has been transformed in reflectance by applying a correction for the solar zenith angle A series of 10 Meteosat visible and infrared images taken at 12h GMT for the 10 day period forms the basis of the data set. The sub-set we consider in this paper is a collection of 60600 pixels, randomly

sampled, from the series of ten meteosat images at 12h GMT. The considered region is an oceanic zone over the Atlantic sea between the tropics.

Many methodologies have already been developed for extracting cloud information from VIS-IR imagery of geostationary satellites, for an historical summary see [8], and[4]. In many cases, for an efficient separation in cloud type and an exact detection of the clear sky zones, the use of VIS/IR spectral signature is not sufficient. From a top of the atmosphere observation, clear sky, broken clouds or semi-transparent clouds can be almost equivalent from a simple point vis-ir measurement at 1 to 10km scale. This is also the case for middle thick clouds and thin cirrus above lower thick clouds. To resolve the ambiguities, the importance of information on the spatial variability of radiance field such as the simple local IR variance, has been stressed by [3]. In this experiment we have considered four feature for describing each pixel: IR radiance, VIS radiance, IR local variance and VIS local variance (computed over a 5×5 sliding window).

3.1 Results



(a) k-means Clustering

(b) Hierarchical Clustering

4 Conclusion

The neural network based method we have described in this paper works well compare to classical k-means and produce automatically coherent results. The approach makes it very easy to have different level of classification....

References

- [1] C. Ambroise. *Approche probabiliste en classification automatique et contraintes de voisinage*. PhD thesis, Université de Technologie de Compiègne, 1996.
- [2] C. Ambroise and G. Govaert. Constrained clustering and kohonen self-organizing maps. *Journal of Classification*, 13(2):299–313, 1996.
- [3] J. A. Coakley and F.P. Bretherton. Cloud cover from high-resolution scanner data: Detecting and allowing for partially filled fields of view. *J Geophys. Res.*, 87:4917–4932, 1982.
- [4] A.H. Goodman and A. Henderson-Sellers. Cloud detection and analysis: a review of recent progress. *Atmospheric Research*, 21:203–228, 1988.
- [5] S.P. Luttrell. Derivation of a class of training algorithms. *IEEE transactions on Neural Networks*, 1(2), 1990.
- [6] F. Murtagh. Interpreting the kohonen self-organization feature map using contiguity-constrained clustering. *Pattern Recognition Letters*, 16:399–408, 1995.
- [7] B.D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge university press, 1996.
- [8] W.B. Rossow, C.L. Brest, and L.C. Garder. Global seasonal surface variations from satellite radiance measurements. *Journal of Climate*, 2:214–247, 1989.
- [9] S. Thiria, Y. Lechevallier, O. Gascuel, and S. Canu. *Statistique et méthodes neuronales*. Dunod, 1997.

ACKNOWLEDGEMENTS

This research was supported by the NEUROSAT european research project (PL952049).