Supervised Classification Using Gaussian Mixture Models

Ana Claudia Oliveira de Melo, Ronei Marcos de Moraes and Liliane dos Santos Machado

UFPB–Universidade Federal da Paraíba - Cidade Universitária s/n, João Pessoa, PB, Brasil
{anaclaudia,ronei}@de.ufpb.br, liliane@di.ufpb.br

Abstract. This paper proposes the use of Gaussian Mixture Models as a supervised classifier for remote sensing images. We present some results of this method application over a real image.

1 Introduction

Several times, researchers in image processing justify the use of Gaussian approach due to data volume. In the practice of image classification, this approach should not be applied without verify the Gaussian distribution hypothesis. It could implicates in a low percentage of correct classification, as example, when the Maximum Likelihood classifier is utilized. The Gaussian Mixture Models (GMM), in theory, can be applied to modeling a large number of distributions, including non-symmetrical distributions. These models have been used to data classification and speech and speaker recognition.

2 The GMM classifier

In image processing applications, researchers commonly use the unsupervised version of mixture models [1, 3]. In this paper, we proposed the utilization of ellipsoidal GMM [4] as a supervised classifier for remote sensing images. In the GMM each x_i is assumed to arise independently from a mixture with density $f(x_i|\theta) = \sum_{k=1}^{K} p_k h(x_i,|\mu_k,\Sigma_k)$, where p_k is the mixing proportion ($0 < p_k < 1$ for all $k = 1, ..., K$ and $p_1 + ... + p_k = 1$) of k^{th} component and $h(\bullet|\mu_k,\Sigma_k)$ denotes the d-dimensional Gaussian density with mean μ_k and variances matrix Σ_k. Thus, the vector of the mixture parameters is $\theta = (p_1, ..., p_K, \mu_1, ..., \mu_K, \Sigma_1, ..., \Sigma_K)$. The clusters associated to the mixture components are ellipsoidal, centered at the means μ_k and variances matrix Σ_k determine geometric characteristics of the ellipses. To estimate the parameters for the model, we can use the classical Expectation-Maximization (EM) algorithm [2], which has as goal to find the maximum likelihood estimators for θ and that maximize the observed log-likelihood function in m interactions. To evaluate the quality of estimations, we use Bayesian Information Criterion (BIC) [4], which compares models with different parameterizations or different number of components. To test the classifier proposed, we performed a case study with a real image of Brazilian state of Pará at Tapajós River. The image was obtained from the number 4 spectral sensor of the Landsat satellite. The image was composed by three classes: Tapajós River (class 1), Contact Areas (class 2) and Human Occupation Areas (class 3). All classes were modeled using two components ($K = 2$), because those distributions were not Gaussian distributions. The BIC confirms the quality of models.

A statistical comparison between classification and a reference map of the same area was performed using Kappa coefficient. The Kappa results were 88.6875% with variance 1.1416×10^{-4}. The same coefficient for each class results were: 99.6929% for class 1, 78.8462% for class 2 and 86.4718% for class 3. The main misclassification occurs with classes 2 and 3. In fact, there is an overlapping of those classes' distributions.

Thus, the classification can be considered satisfactory in statistical terms, even if we considerer the overlapping of class 2 and class 3 distributions. It is important to note that these results were obtained using only one original spectral band. As future works we intent to do a performance statistical comparison with others statistical classifiers.

References

[1] H. Caillol, W. Pieczynski and A. Hillon, Estimation of fuzzy Gaussian mixture and unsupervised statistical image segmentation, *IEEE Trans. on Image Proces.*, 3(6), 425-440, 1997.

[2] A.P. Dempster, N.M. Laird and D.B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm. *J.R.Statis.Soc.*,B,39, 1-38, 1977.

[3] W. Pieczynski, J. Bouvrais, and C. Michel, Estimation of generalized mixture in the case of correlated sensors, *IEEE Trans. on Image Proces.*, 2(9), 308-311, 2000.

[4] G. Schwartz, Estimating the Dimension of a Model. *The Ann. of Stat.*, 6, 461-464, 1978.