Co-clustering for hyperspectral images.

Julien Jacques, Cyril Ruckebusch

COCLUSTERING FOR HYPERSPECTRAL IMAGES

J. Jacques [1], C. Ruckebusch [2]

[1] *ERIC , Université de Lyon, Lumière, Bron, France*
[2]*LASIR, Université Lille 1, Villeneuve d'Ascq, France*
julien.jacques@univ-lyon2.fr

Keywords (3-5) : Co-clustering, Guassian models, Spatial-spectral, Hyperspectral imaging

Clustering is often used for hyperspectral images in order to assign sets of pixels into a number of different homogeneous groups called clusters. As a result, pixels in the same cluster have similar spectra, i.e. are close to each other in a certain sense. Clustering is a core technique of the chemometrics toolbox but some limitations can be pointed for hyperspectral imaging. A first limitation of clustering is that it only considers information in the spectral dimension. Another is that it groups whole vectors. This means that if one or a few elements of the vectors differ significantly, the vectors cannot be clustered together. These limitations may result in suboptimal grouping.

Co-clustering [1] is a generalization of unsupervised clustering. Co-clustering approaches allow finding blocks of similar data in a matrix by considering simultaneously information along the rows and columns. It has drawn attention in chemometrics [2] and, more recently, in hyperspectral imaging [3]. Applied in hyperspectral image analysis, co-clustering is able to simultaneously group the pixels and the spectral features, building co-cluster (blocks) incorporating both spatial and spectral information and revealing patterns/relationships between them.

In this work, we propose using a co-clustering technique based on Gaussian mixture models for high-dimensional data [4]. This appraoch for co-clustering consists in assuming that the distribution of the observations into a block (combination of pixels and wavelengths) have a common univariate distribution, that we can for instance be assumed to be Gaussian for hyperspectral images. Model parameter estimation in the co-clustering context is not as straightforward as in clustering. Indeed, the traditional EM algorithm is not computationally tractable since all the possible blocks cannot be enumerate in the traditional E step. To overcome this problem, we consider a Stochastic EM algorithm incorporating a Gibbs sampler into the SE step in order to simulate the row and column partitions. A further step would be to incorporate information of spatial proximity of pixels and of spectra proximity of wavelength within the Gaussian latent block model. The main challenge is to define a co-clustering model which could be able to take into account these proximities, in order to exhibit compact clusters. In this talk we will show the results obtained on hyperspectral Raman imaging data and remote sensing data.

References:

[1] G. Govaert, M. Nadif, *Co-Clustering*, Wiley (2013).
[2] Bro et al. *Journal of Chemometrics* 26 256-263 (2012).
[3] Emma Izquierdo-Verdiguier, *IEEE International Geoscience and Remote Sensing Symposium* 465-468 (2015).
[4] Jacques J., Bouveyron C., Girard S., Devos O., Duponchel L., Ruckebusch C. Gaussian mixture models for the classification of high-dimensional vibrational spectroscopy data, Journal of Chemometrics 24 [11-12], 719-727 (2010).