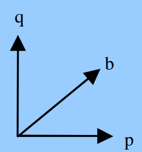
Reduction and segmentation of hyperspectral data cubes

Astronomical Data Analysis Software & Systems XVI
15-18 October 2006

F. Genova [2], M. Petremand [1,2], F. Bonnarel [2], M. Louys [1], Ch. Collet [1]

- *Considering an hyperspectral data cube as an image of spectra, the purpose of this work is to extract common spectral behaviors from the data and combine both spectral and spatial information. Such analysis on data cubes, whose size increases drastically (typically 1000x1000x1000 sampling elements) faces the curse of dimensionality. Therefore we propose a prior step as an iterative dimensionality procedure based on the Meanshift approach.*
- *A priori knowledge → define a reference basis of spectra, seen as 1D vectors. After projecting each spectrum on this basis, a set of projection weights is built up. The Meanshift algorithm is used to find the modes of the projection weights. These modes in turn lead to a new projection basis and the process can iterate . The same approach has been tested successfully using a vector angle instead of projection weight.*
- This new method is under validation on simulated galaxy fields → GALICS (Hierarchical Galaxy Formation - IAP)

Hyperspectral Image



- The hyperspectral observation Y is composed of $p \times q \times b$ pixels. Each spatial location s defines a spectrum Y_s.
- n spectra randomly chosen in the observation define the first projection basis B.
- When the spectral behaviors present in the observation are known, the projection basis can thus be composed of reference spectra selected among a spectrum database.

Y_s $\quad B$

Projection or angle of each spectrum on a reference basis

- $B = \{\beta_1 \cdots \beta_n\}$ is the projection basis. Y_s is projected on B
- We can then write the reconstruction $\hat{Y}_s$ of Y_s :

$$\hat{Y}_s = \sum_{i=1}^{n} \omega_{si} \times \beta_i$$

- The weights ω_{si} for each Y_s of the image define a point distribution in the projection space

- As the projection is luminosity variant, one can also use a spectral angle measure :

$$\alpha(\beta_i, Y_s) = arccos\left(\frac{<\beta_i, Y_s>}{|\beta_i||Y_s|}\right)$$

- This measure is luminosity invariant and efficient to highlight spectral behaviors

ω_{si}

New projection basis

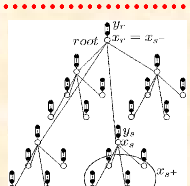
Meanshift

modes

number of classes $\quad \omega_{si}$

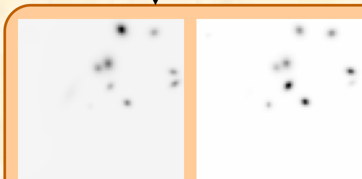
- Meanshift : a non-parametric method to find the modes (maxima) of a probability density function f of a points-distribution.
This method is based on the estimation of the gradient of f : $\nabla f(x)$
- Used in the projection/angle space for each weight/angle vector → Each vector corresponds to a mode.
- The spectra associated to their modes in the data cube produce a new projection basis. Projection and meanshift steps are done again until convergence (same basis during two steps).
- These modes are characteristic of the main spectral behaviour present in the image → number of spectral classes.

Multicomponents segmentation in the projection space : Kmeans or Markovian segmentation

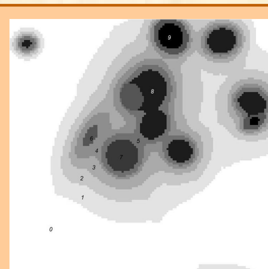


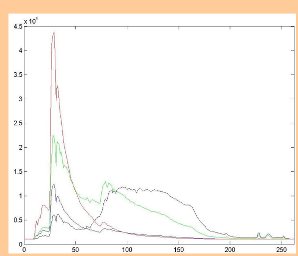
Quadtree structure

- When $n < 10$ a quadtree-based Markovian segmentation leading to a segmentation map X is carried out in the projection or angle spaces. This Markovian segmentation introduces a spatial regularization and clusters weigths/angles into homogeneous regions.
- When $n > 10$ a K-means clustering algorithm is used for the segmentation. Indeed the Markovian model can't be carried out when $n > 10$ because of the "curse of dimensionnality".
- The spectrum number into the last basis gives access to an estimation of the class number for the segmentation step. Moreover the users can set the class number themselves.

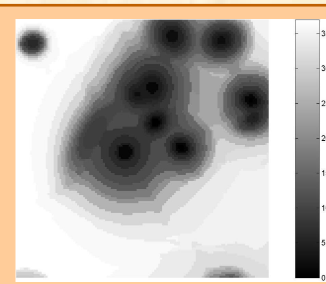


2 bands from the original hyperspectral cube ($128 \times 128 \times 262$ pixels). Object luminosity radially decreases. Each object has a flux variation according to the wavelengths





Segmentation map (10 classes) obtained with the angle method. The spectral angle measure only extracts spectral classes. At the center, the class *bulge* dominates whereas at the periphery the class *disk* is present. The mean spectra for classes 7 (blue), 8 (black), 9 (green), 10 (red) are shown to the right of the segmentation map.



Segmentation map (38 classes) obtained with the projection method. Within an object, the class set is composed of a luminosity class sequence and a bulge-disk spectral composition.

This research is supported by the French government (ACI MDA project: Masses de Donnes en Astronomie http://cdsweb.u-strasbg.fr/MDA/mda.html)

petremand@lsiit.u-strasbg.fr

(1) Laboratoire des Sciences de l 'Image, de l 'Informatique et de la Télédétection, UMR 7005 CNRS, University of Strasbourg I, FRANCE
(2) Observatoire de Strasbourg, CDS, UMR 7004 CNRS, University of Strasbourg I, FRANCE