Binary Partition Tree for hyperspectral remote sensing images

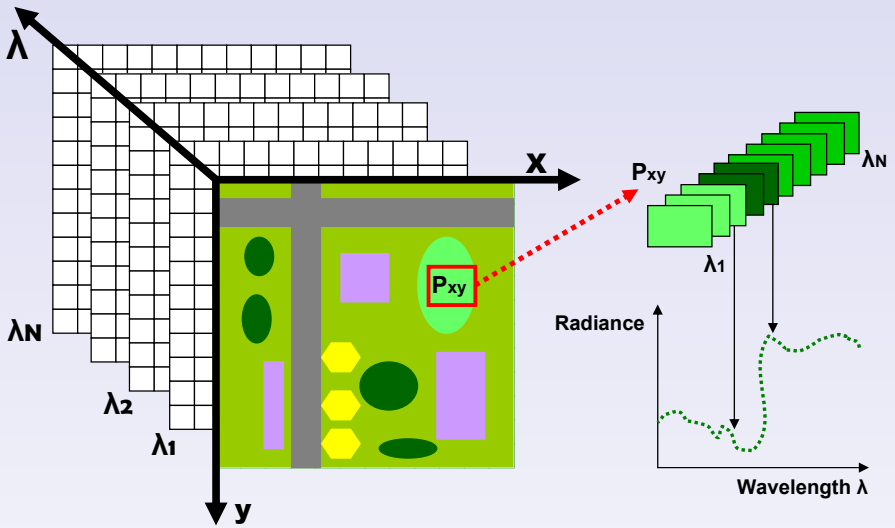
S.Valero [1,2], P.Salembier[2] and J.Chanussot[1]

[1]**GIPSA-lab, Signal & Image Dept., Grenoble Institute of Technology, Grenoble, France.**
[2]**Signal Theory and Communications Department, Technical University of Catalonia, Spain**

Outline

Hyperspectral Imagery

Hyperspectral data cubes contain hundreds of contiguous waveband images



Each pixel is a discrete spectrum containing the reflected radiance of the spatial region that it represents

Different analysis techniques have been proposed in the literature studying these images from their pixel representation.
Images are studied as an array of spectral data without any spatial structure

Hyperspectral Imagery

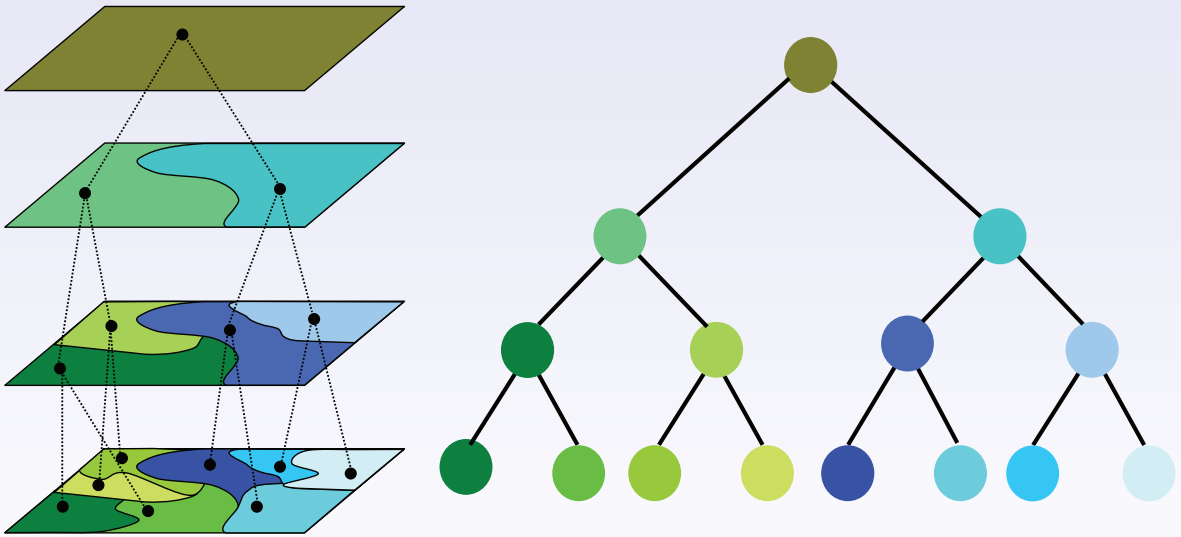
- The initial pixel-based representation is a very low level and unstructured representation. This implies that classification, segmentation or detection techniques are not very robust
- Instead of working with a pixel-based representation, a more advanced strategy consists in studying an image using region-based representations.
- One example of such representation is *Binary Partition Trees*[1]. How can they be extended to the case of hyperspectral data ?

[1] P. Salembier and L. Garrido, *Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval* IEEE Trans. Image Processing, 2000

Outline

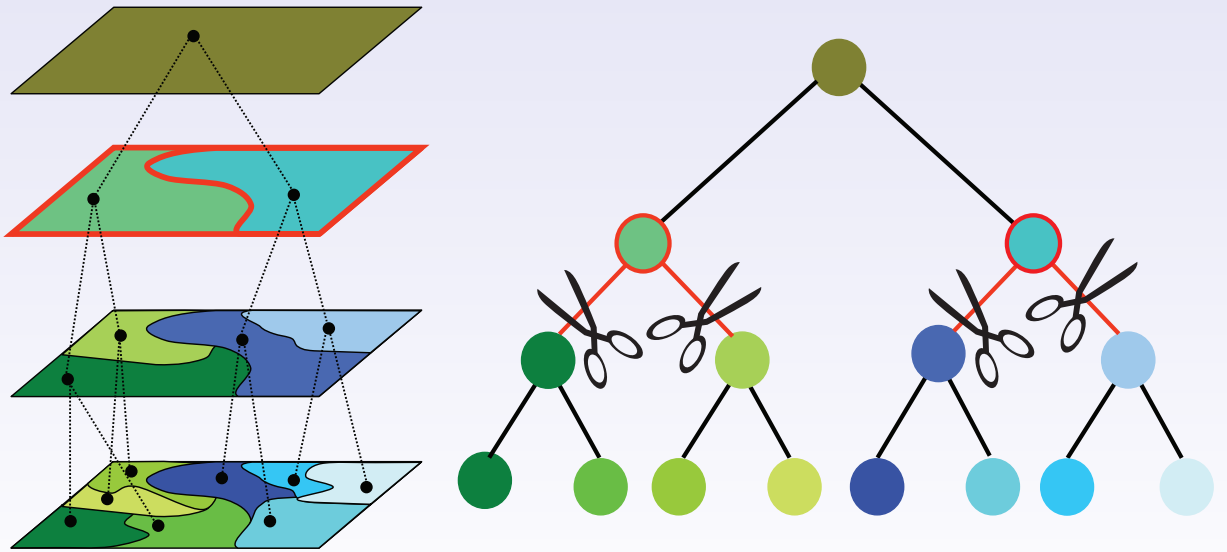
Binary Partition Tree

- BPTs is a set of hierarchical regions stored in a tree structure
- Each node representing a region in the image, BPTs allow us to extract many different partitions



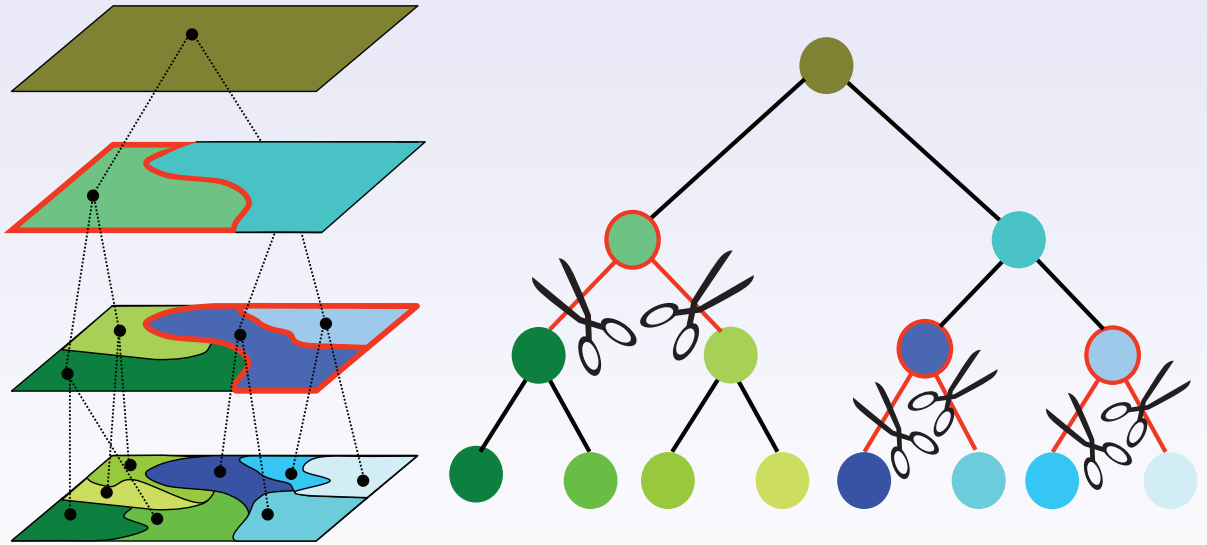
Binary Partition Tree

- BPTs is a set of hierarchical regions stored in a tree structure
- Each node representing a region in the image, BPTs allow us to extract many different partitions

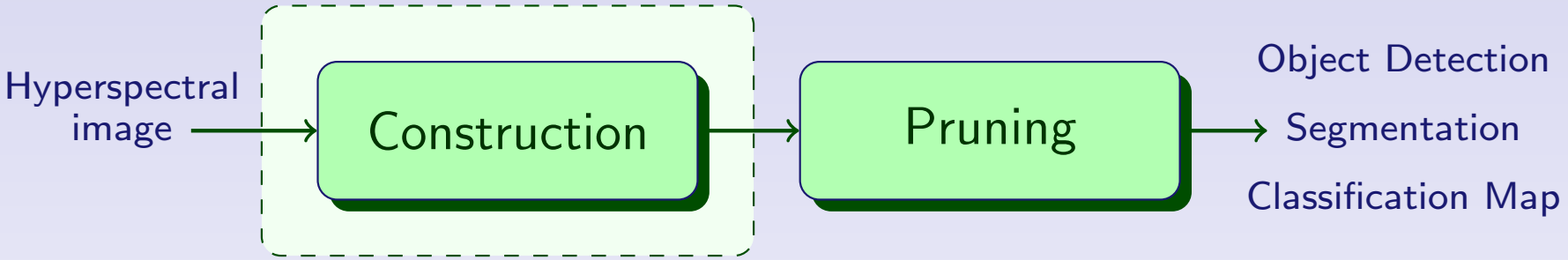


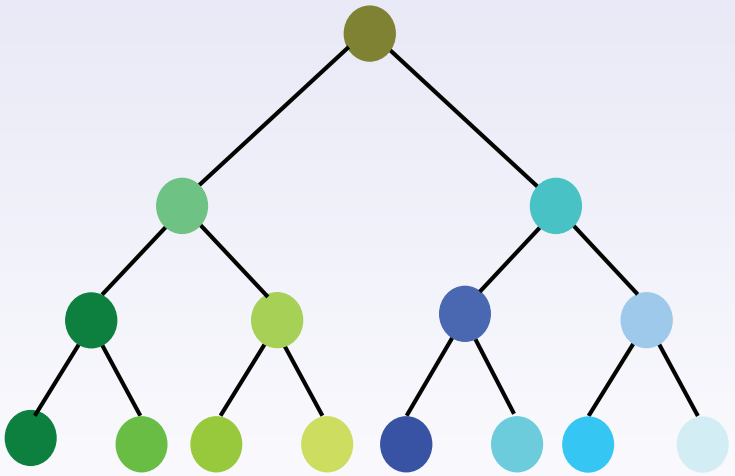
Binary Partition Tree

- BPTs is a set of hierarchical regions stored in a tree structure
- Each node representing a region in the image, BPTs allow us to extract many different partitions



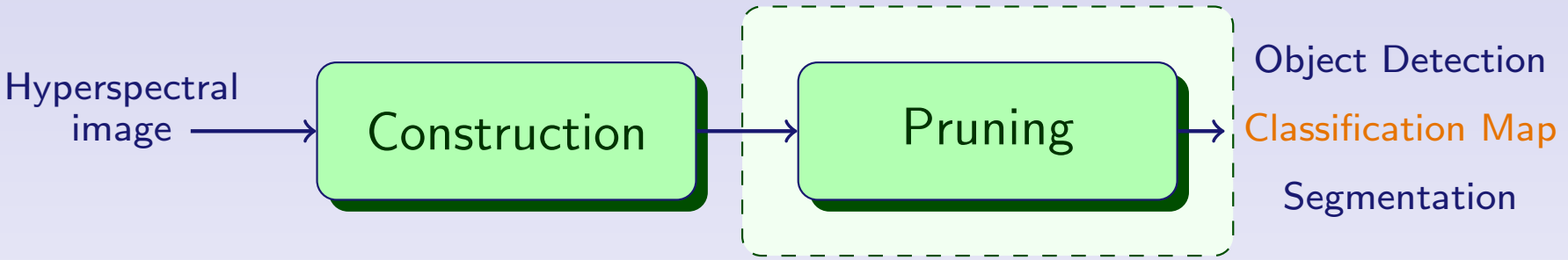
Aim: BPT for HS image analysis

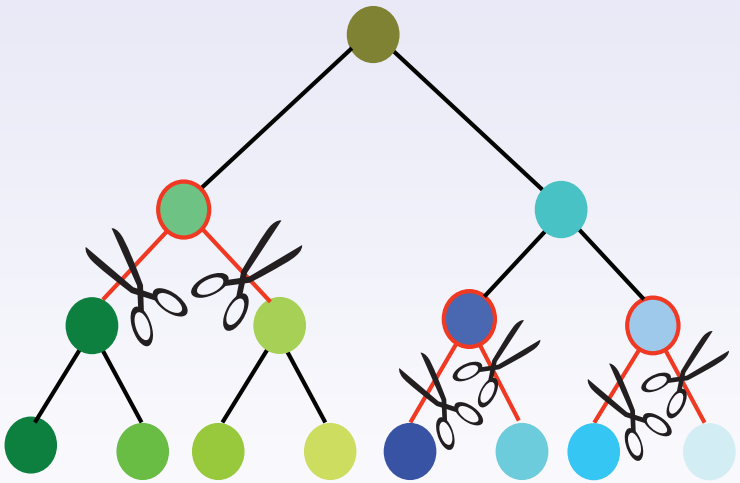




First goal: The construcion of a region-based representation for hyperspectral data

Aim: BPT for HS image analysis



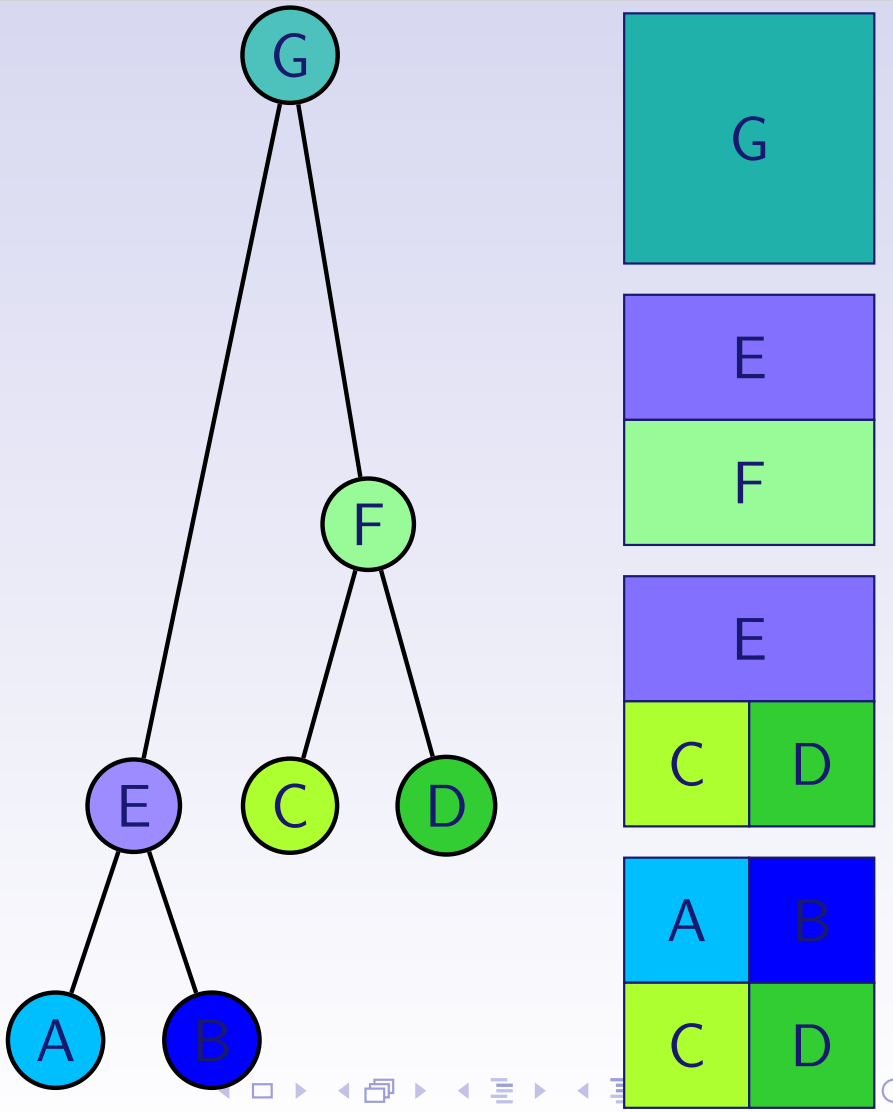


Second goal: Definition of application dependent pruning strategies

Outline

BPT Construction

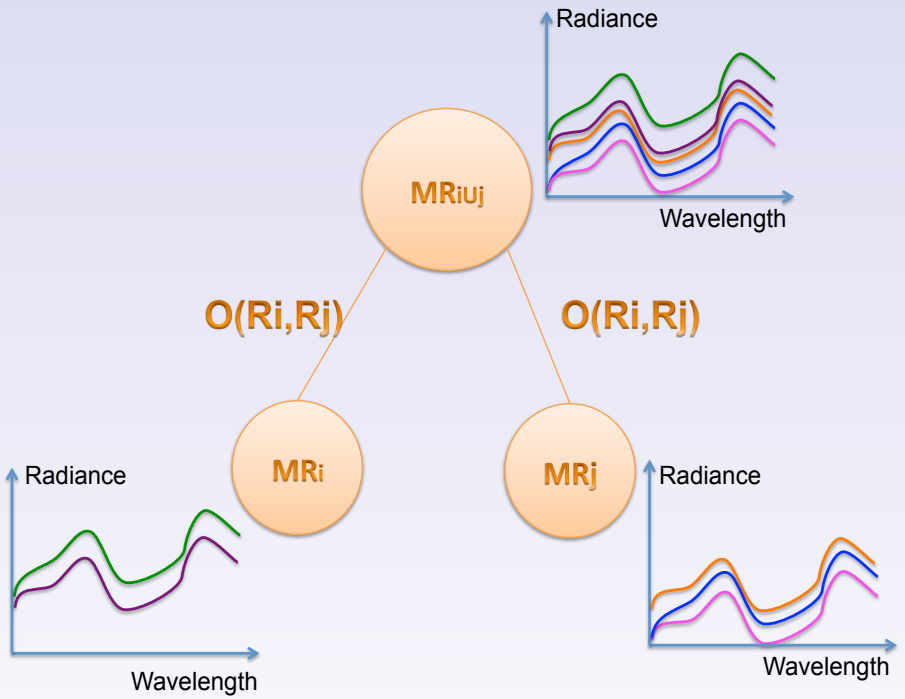
- The BPT is a hierarchical tree structure representing an image
- Leaves: Image pixel
  Root: entire image support
- Remaining node: region formed by the merging of 2 children
- Tree construction: iterative region merging



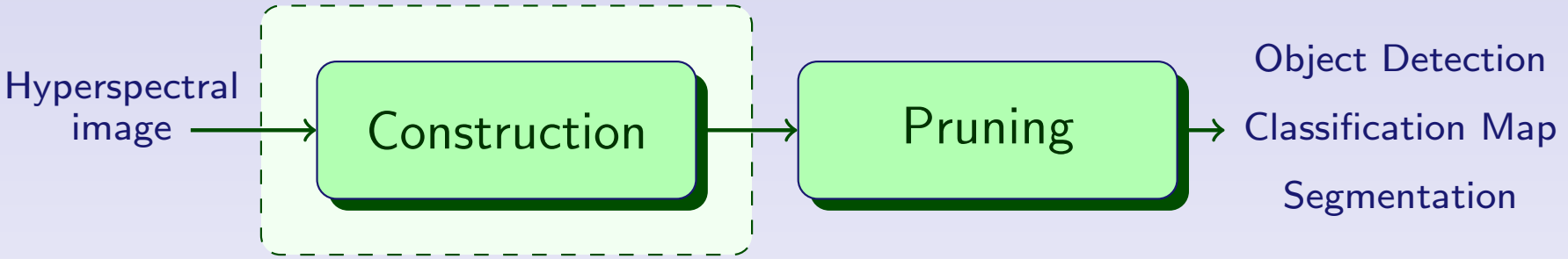
BPT Construction

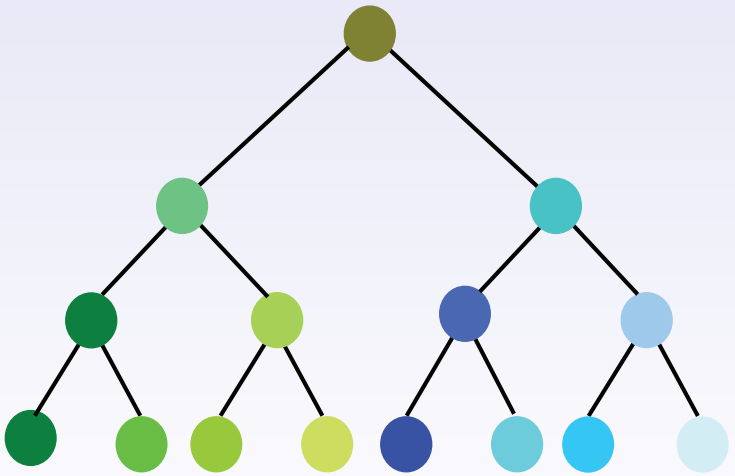
The creation of BPT implies two important notions

- Region model M_{R_i}
  How to represent hyperspectral region?
  How to represent the union of 2 regions ?
- Merging criterion $0(R_i, R_j)$
  Which pair of regions have to be merged?



Merging Order Definition



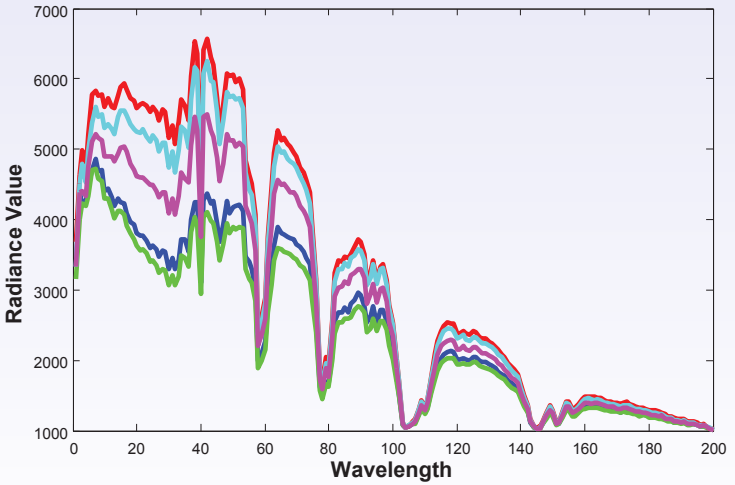


- How to represent hyperspectral image regions ?
- Which similarity measure defines a good merging order ?

Outline

Region Model

- Due to factors such as the noise from atmospheric conditions, the sensor influence or the illumination effect,a large variability exists inside a set of spectra corresponding to the same object
- Given a set of spectra corresponding to the same object, the homogeneity of the radiance values cannot be assumed in each wavelength

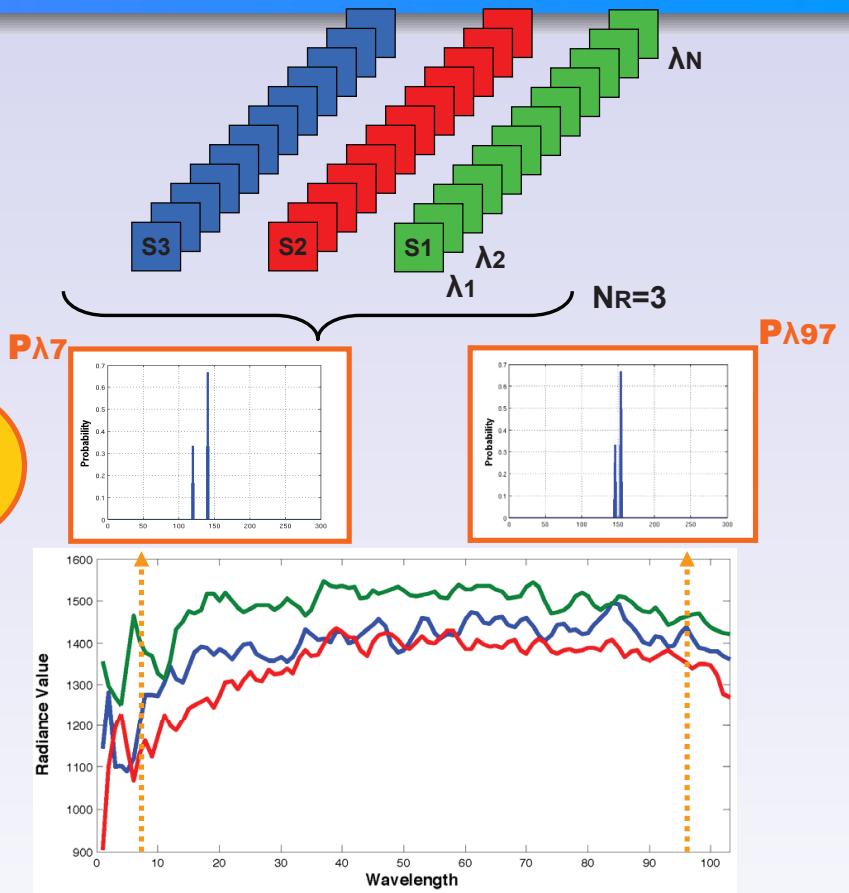


Non Parametric statistical Model

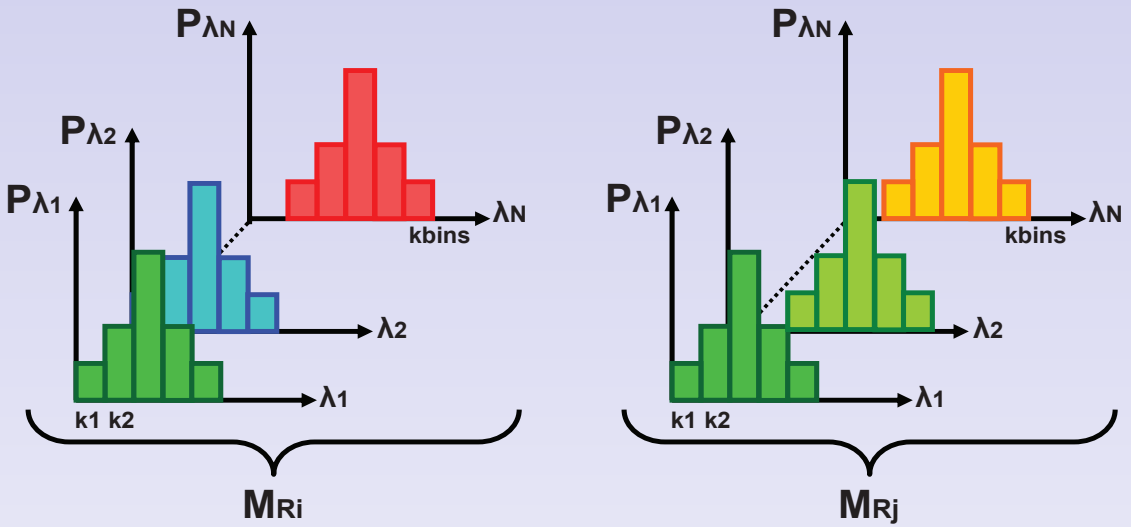
Non-parametric statistical region model *[2]* consisting in a set of N probability density functions

$$M_R = \{P_{\lambda 1}, P_{\lambda 2}, ..., P_{\lambda N}\}$$

where each $P_{\lambda i}$ represents the probability that the spectra data set has a specific radiance value in the wavelength λi.



[2] F. Calderero and F. Marques, *Region Merging Techniques Using Information Theory Statistical Measures*,IEEE Transactions on Image Processing, 2010



$$M_{Ri \cup Rj} = \{P_{\lambda 1}^{Ri \cup Rj}, P_{\lambda 2}^{Ri \cup Rj}, ..., P_{\lambda N}^{Ri \cup Rj}\}$$

where $P_{\lambda z}^{Ri \cup Rj}$ of the union of two regions can be defined as

$$P_{\lambda z}^{Ri \cup Rj}(k) = \frac{N_{Ri}}{N_{Ri}+N_{Rj}} P_{\lambda z}^{Ri}(k) + \frac{N_{Rj}}{N_{Ri}+N_{Rj}} P_{\lambda z}^{Rj}(k)$$

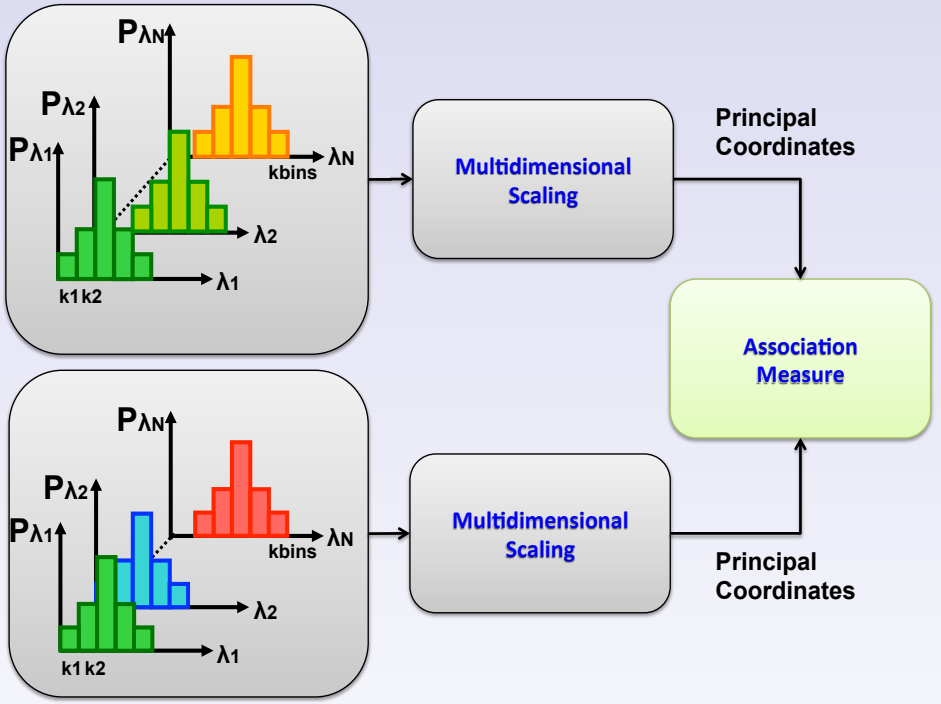
being N_{Ri} and N_{Rj} the area of the regions Ri and Rj.

Outline

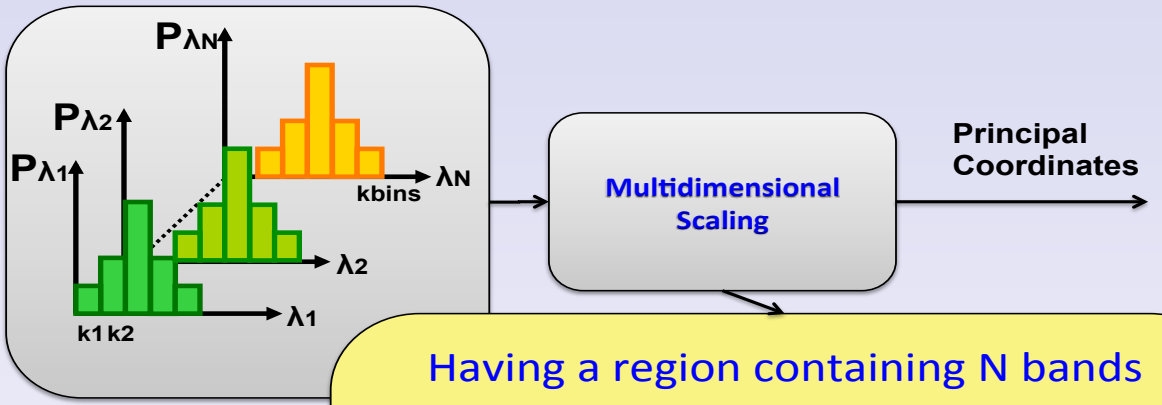
Merging Criterion

The merging criterion consists in

- To analyze the inter-waveband similarity relationships for each data via metric scaling and principal coordinates
- To establish a similarity measure correlating the principal axis of both data sets



Merging Criterion: MDS



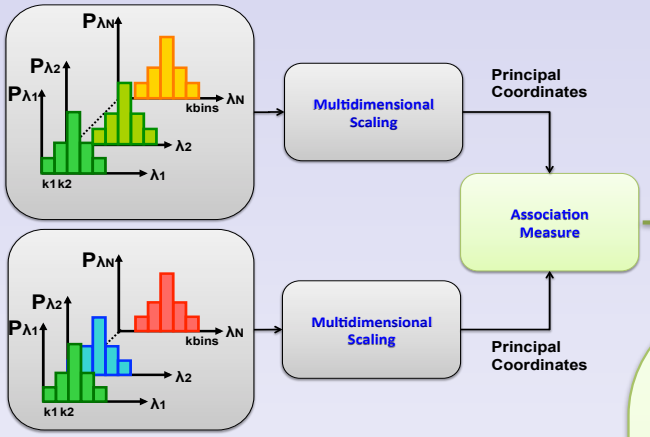
Having a region containing N bands

1) We construct a NxN distance matrix representing the distance between all bands **

2) We decompose this matrix to find the first principal coordinates

[1] H. Ling and K. Okada *Diffusion distance for histogram comparison*. In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2006

Merging Criterion



Given X and Y, two data sets representing the principal coordinates of two hyperspectral regions

An asociation measure is defined by considering that X and Y are the predictor and the reponse variable of a multivariate regression model

$Y = XB + e$

Merging Criterion

Being the multivariate linear regression model $Y = X\beta + e$

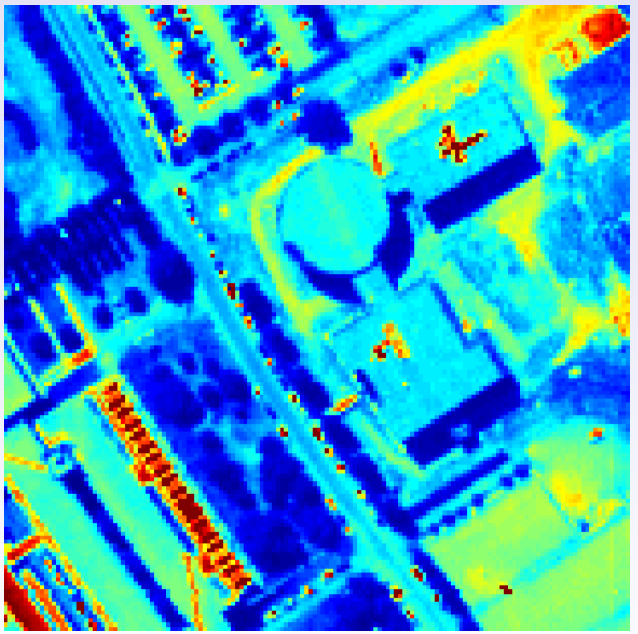
- β is the matrix containing the regression coefficients and e is the error matrix
- Clearly, if there is no relationship between X and Y, the matrix β is equal to 0
- The idea is to compute a Lamba Wilks test veryfing $\beta = 0$ to measure if there is a significant relationship between X and Y

[1] C.M Cuadras, S.Valero, P.Salembier and J. Chanussot *Distance-based measure of association with applications in relating hyperspectral images*. In Communications is Statistics, 2010

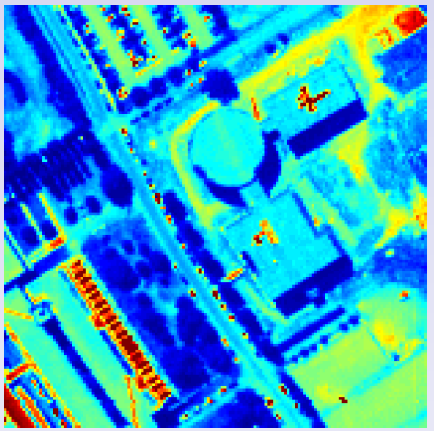
Outline

ROSIS Pavia University

- Hyperspectral ROSIS Pavia University, containing 103 bands
- An image band is shown in the image below



Construction Example



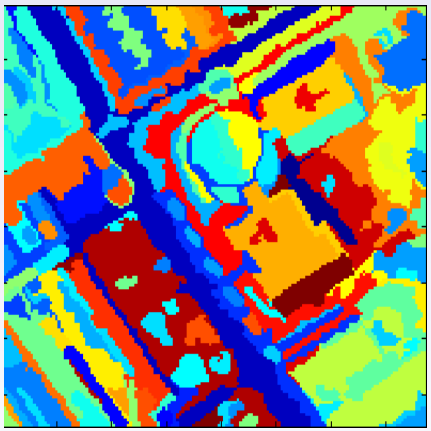


Figure: 150 regions

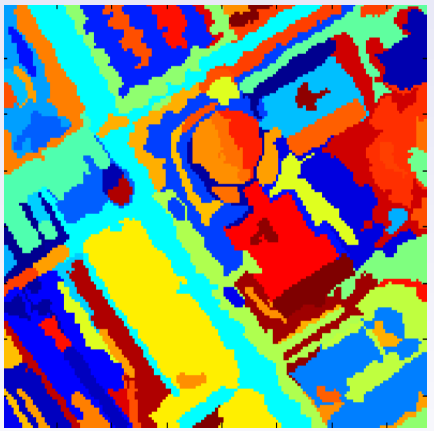


Figure: 100 regions

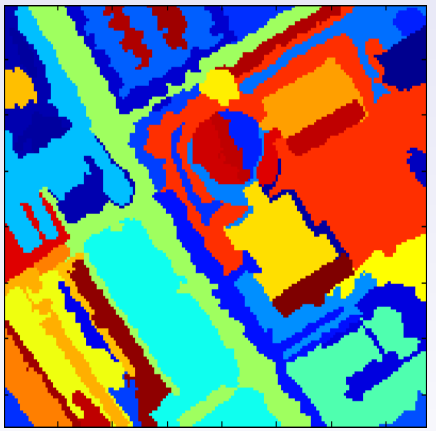


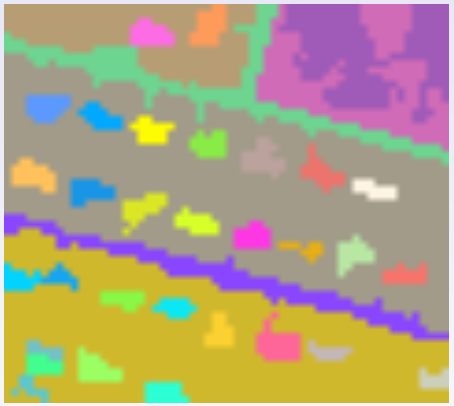
Figure: 50 regions

HYDICE Urban Area

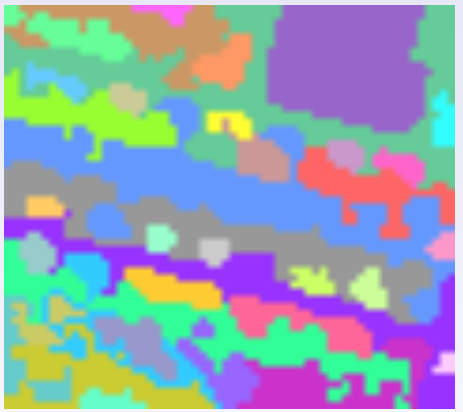
- Hyperspectral HYDICE Urban, containing 167 bands
- An RGB composition image is shown in the image below

BPT Construction Results

RGB Composition
Hydice, Urban image
167 wavebands



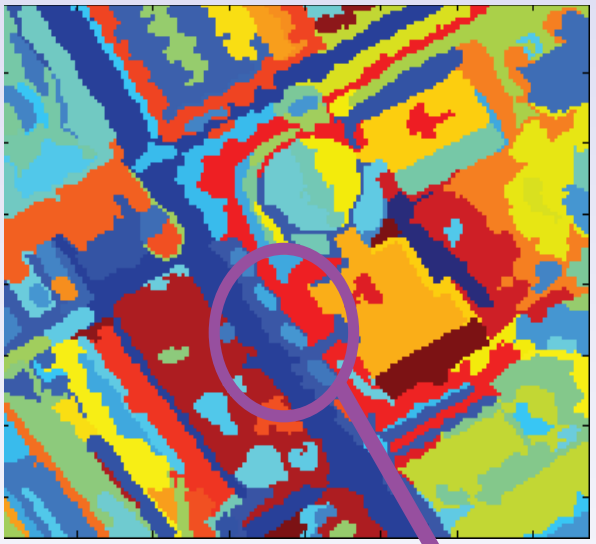
Partition obtained by BPT containing 37 regions



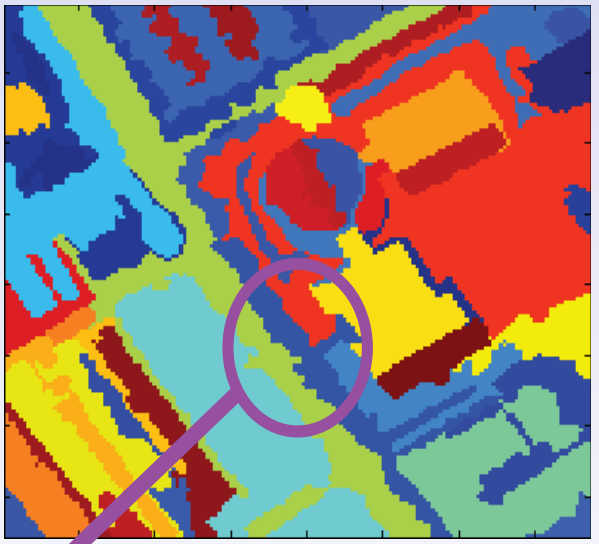
Partition obtained with RHSEG software

Construction Example

- We cannot assume that the optimal partition can be obtained stopping the iterative merging of regions



150 regions

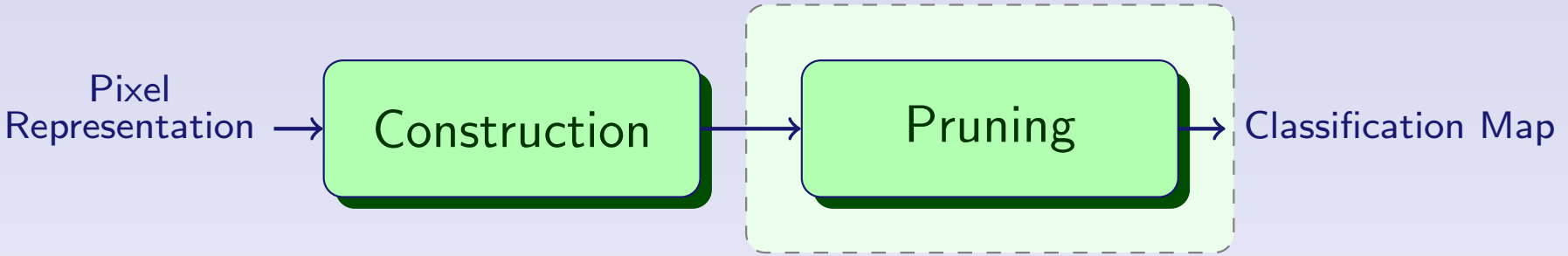


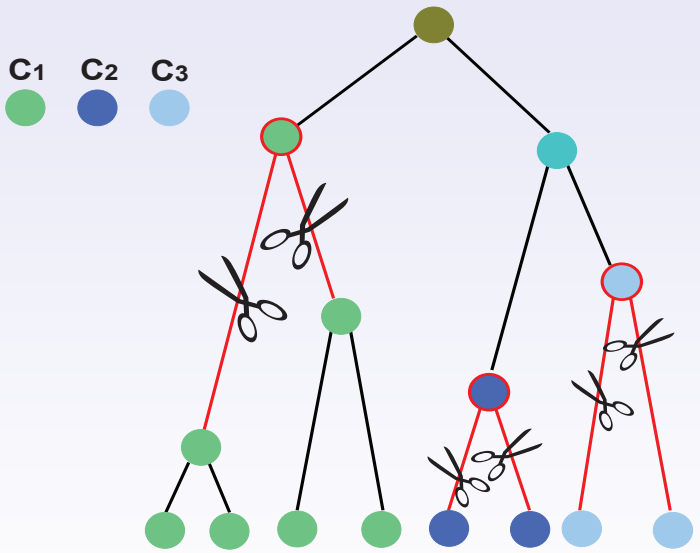
50 regions

Trees have disappeared

Outline

Pruning Strategy



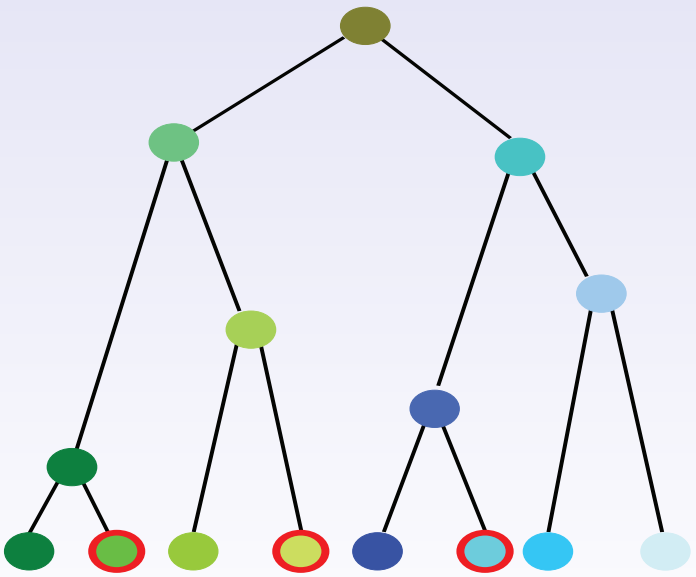


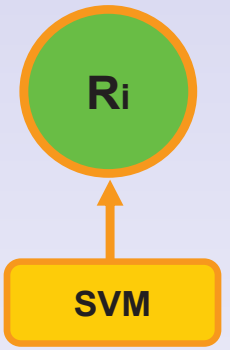
It can be interpreted as a filtering tool that aims to achieve a goal inside a set of hierarchical partition

It consists in removing subtrees composed of nodes belonging to the same class

Classification Pruning

In a classification context, region descriptors can correspond to a classifier output





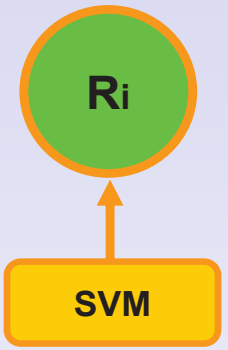
$Class_{Ri}$
$Cp_{Ri} = \{ P_1, P_2, \ldots P_{Nc} \}$

Training the SVM classifier using some leaves nodes

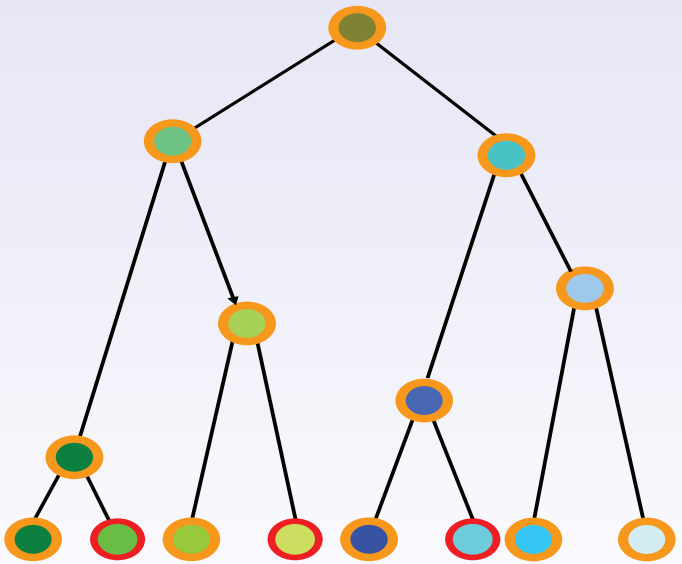
All nodes are populated by their class probability estimation CpR and their predicted class ClassR

Classification Pruning

In a classification context, region descriptors can correspond to a classifier output



$Class_{Ri}$
$Cp_{Ri} = \{ P_1, P_2, \ldots P_{Nc} \}$



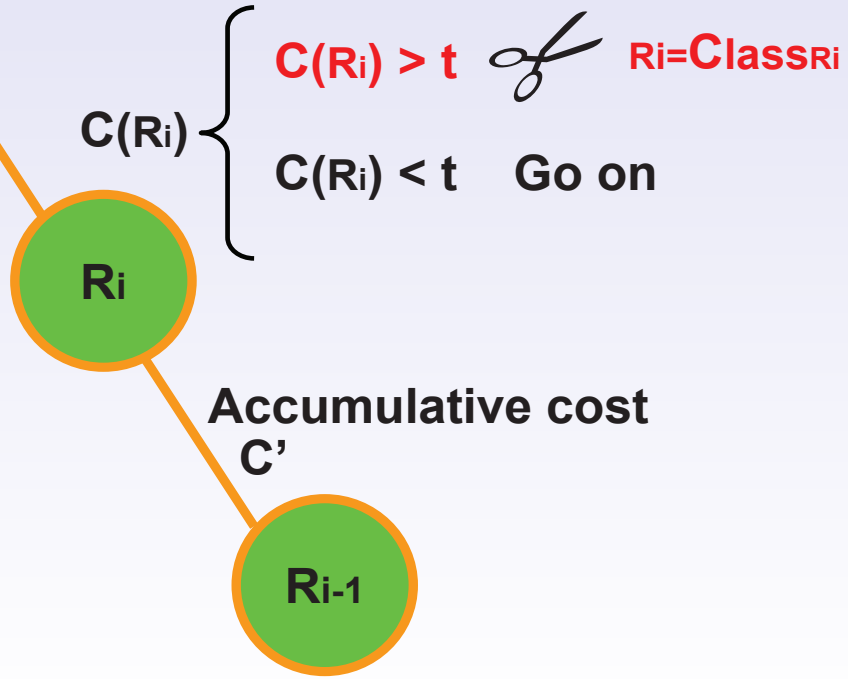
Leave Training Nodes
SVM classified Nodes

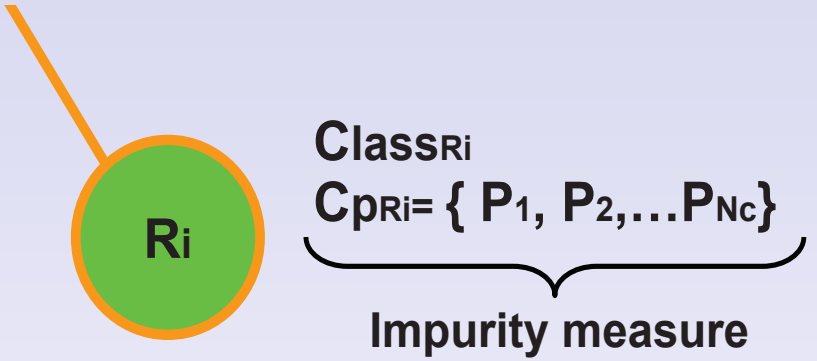
Training the SVM classifier using some leaves nodes

All nodes are populated by their class probability estimation CpR and their predicted class ClassR

Classification Pruning

Classifier output enables to compute an impurity measure for each node

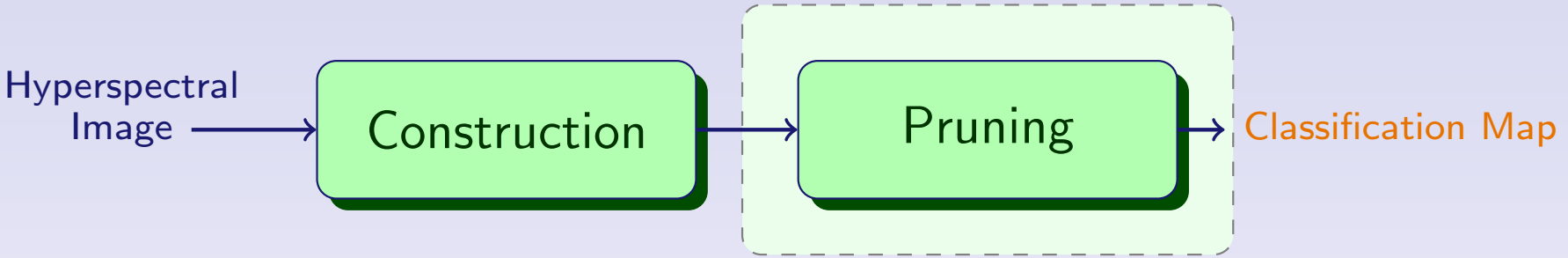


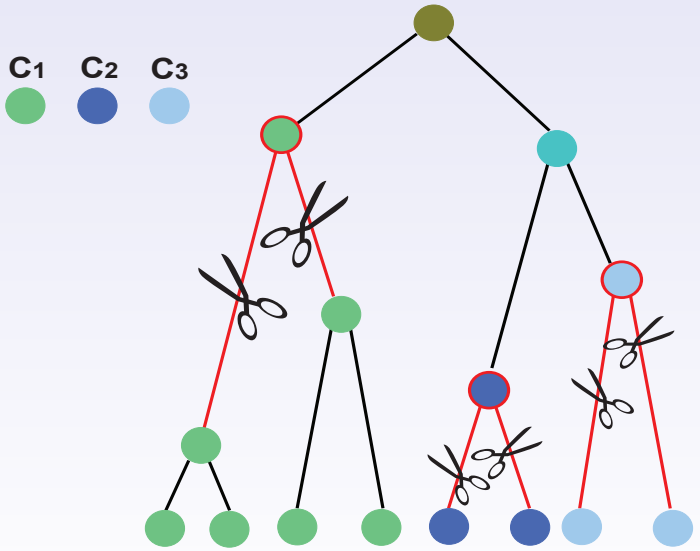


An increasing iterative cost can be studied along the branches.

$$C(R_i) = C' - \sum_{j=0}^{N_c} Cp_{R_i}(j) \log(Cp_{R_i}(j))$$

Experimental Results

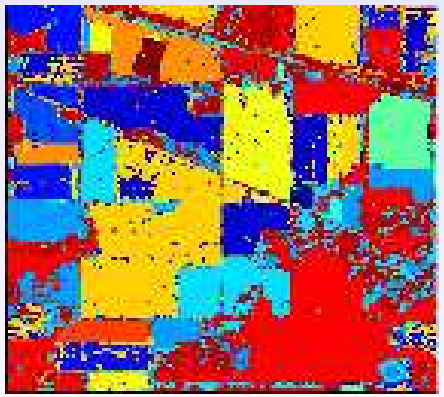




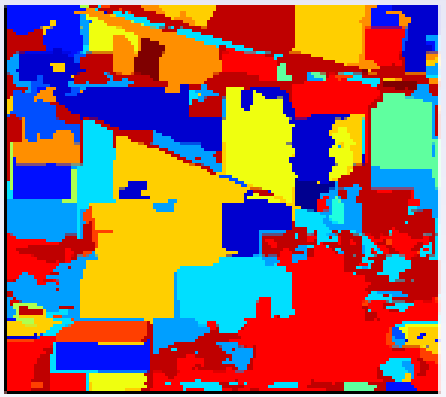
Pruning Example
Hyperspectral Indian Pines
200 wavebands

Classification Pruning

BPT nodes selection according to the proposed pruning criterion provide a less noisy classification



SVM



BPT Pruning

Class	Simple SVM	Pruned BPT
1	86.11	94.44
2	88.39	93.41
3	83.45	89.03
4	77.56	80.77
5	95.18	92.77
6	97.39	98.39
7	83.33	88.89
8	97.85	99.08
9	64.29	100
10	84.98	88.70
11	91.19	96.72
12	92.93	93.66
13	100	98.59
14	96.99	99.42
15	67.72	98.43
16	95.30	100
Overall	87.67	94.52

Outline

Conclusions

- BPTs have been proposed as a new region-based and hierarchical representation of the hyperspectral data
- Being a generic representation, many tree processing techniques can be formulated as pruning strategies for many applications
- A solution for the problem of the spectra variability for clustering hyperspectral data has been proposed using statistical region models

Conclusions

- A new distance-based measure depending on canonical correlations relating principal coordinates have been proposed as similirarity measure

- As a first example of BPT processing, a pruning strategy has been proposed to classify hyperspectral data. Experimental results have shown that the proposed method improves the classification accuracy of a classical SVM