Bayesian analysis in image processing

Florence Tupin
Télécom ParisTech - LTCI

Image classification

- **Introduction**

- **Bayesian classification**
  - Image modeling
  - Mono-spectral case
  - Multi-spectral case
  - Punctual / Contextual

- **Kmeans classification**
  - Unsupervised case
  - algorithm

Introduction

◦ Classification objectives

- identification of the different classes in the image
- preliminary step of pattern recognition methods (object detection)

◦ Hypotheses

- grey-level images
- classes = peaks in the histogram
- low grey-level variations in the same class
- punctual classification : each pixel is classified separatly
- supervised learning : samples of each class are available

◦ Possible extensions

- multi-channel images
- contextual classification : markovian framework

Image classification

◦ **Introduction**

◦ **Bayesian classification**

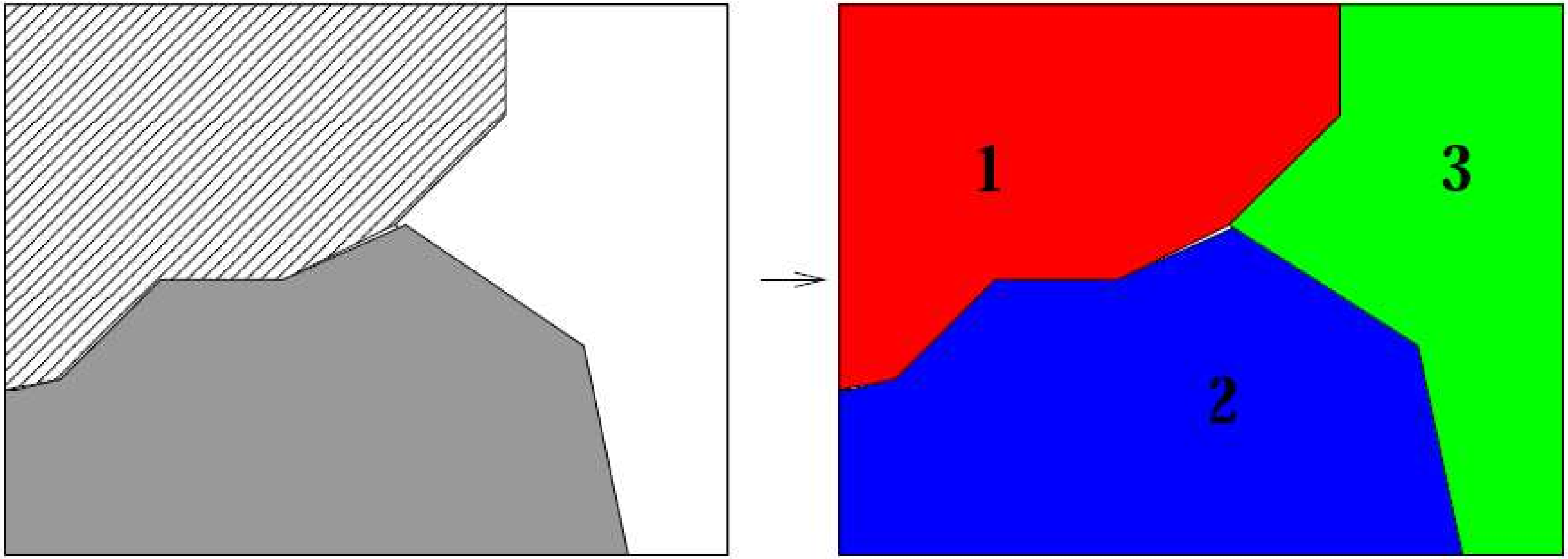
- Image modeling
- Mono-spectral case
- Multi-spectral case
- Punctual / Contextual

◦ **Kmeans classification**

- Unsupervised case
- algorithm

Image modeling

◦ **Probabilistic model**



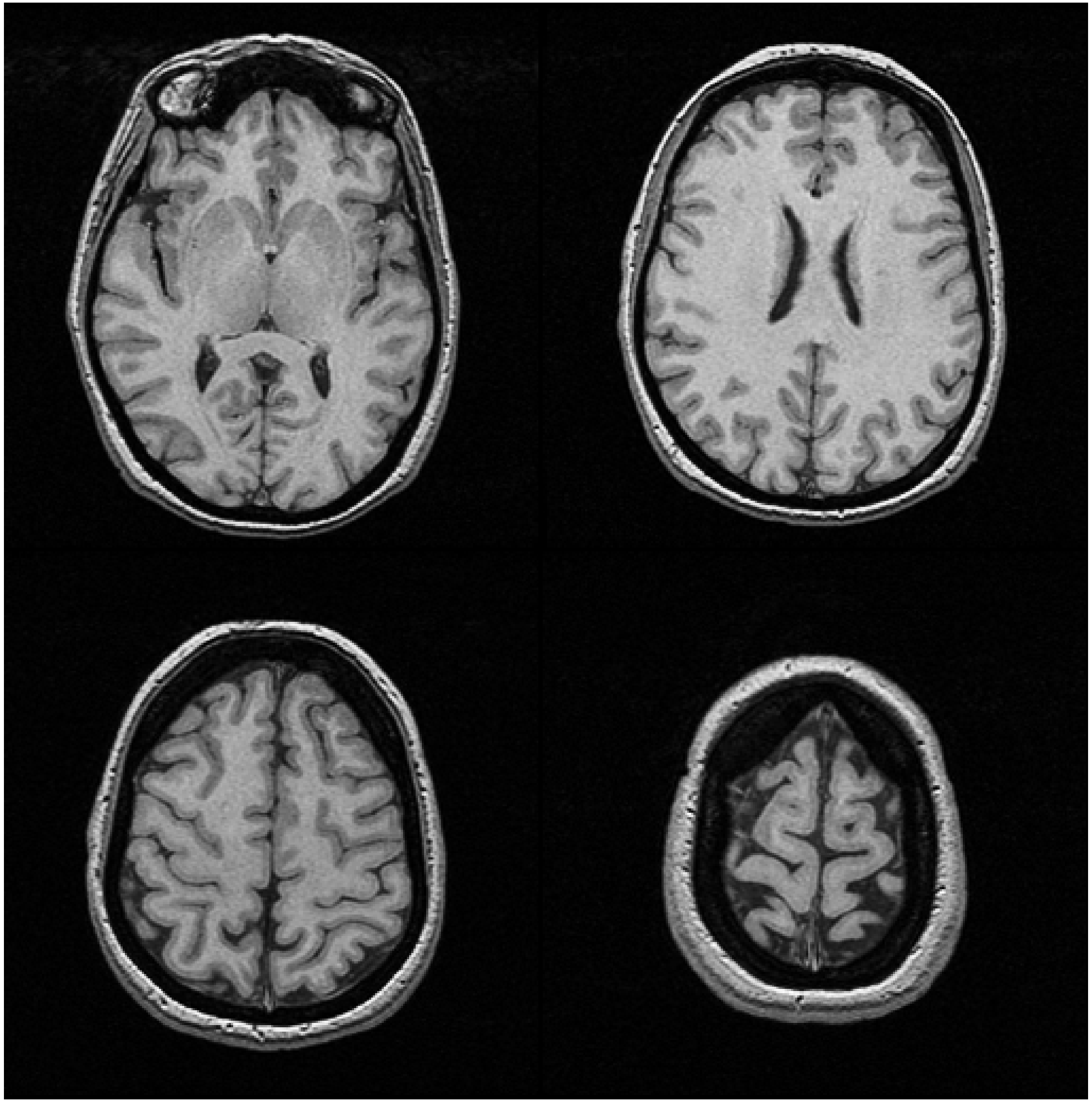
S set of sites (pixels = localization (i, j) in the image)

$$\text{grey-level } y_s \in \{0, ..., 255\} = E \quad \rightarrow \quad \text{class } x_s \in \{1, ..., K\} = \Lambda$$

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

Y_s random variable of grey-level $\qquad$ X_s random variable of label

Example : brain image



Example : brain image

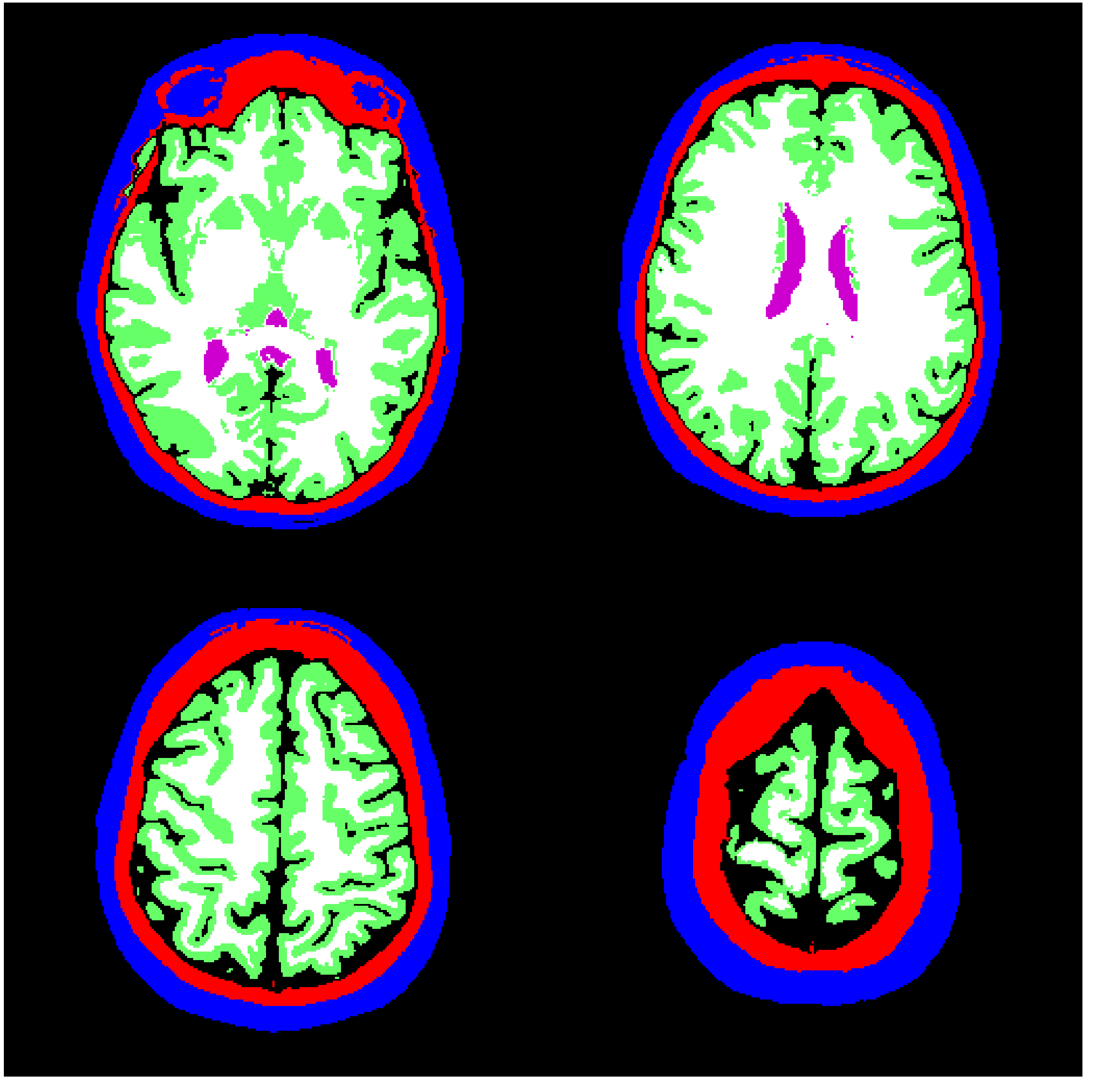


Image classification

◦ **Introduction**

◦ **Bayesian classification**

- Image modeling
- Mono-spectral case
- Multi-spectral case
- Punctual / Contextual

◦ **Kmeans classification**

- Unsupervised case
- algorithm

Mono-spectral case (grey-level image)(1)

- **Maximum A Posteriori criterion**

you know a grey-level y_s for pixel s

$\Rightarrow$ take the “best” class x_s knowing y_s

$\Rightarrow$ find the i which maximizes $P(X_s = i|Y_s = y_s)\ \forall i \in \Lambda$

- **Can we compute $P(X_s = i|Y_s = y_s)$?**

Bayes rule

$$P(X_s = i|Y_s = y_s) = \frac{P(Y_s = y_s|X_s = i)P(X_s = i)}{P(Y_s = y_s)}$$

$$X_s = argmax_{i \in \{1,\ldots,K\}} P(Y_s = y_s|X_s = i)P(X_s = i)$$

Mono-spectral case (grey-level image) (2)

$$P(X_s = i | Y_s = y_s) = \frac{P(Y_s = y_s | X_s = i) P(X_s = i)}{P(Y_s = y_s)}$$

◦ **Example of brain image**

$\Lambda = \{0 = \emptyset; 1 = skin; 2 = bone; 3 = GrayMatter; 4 = WhileMatter; 5 = LCR\}$

• $P(X_s = i)$ = apparition probability of class i

• $P(Y_s = y_s | X_s = i)$ = grey-level distribution knowing that the pixels belong to class i

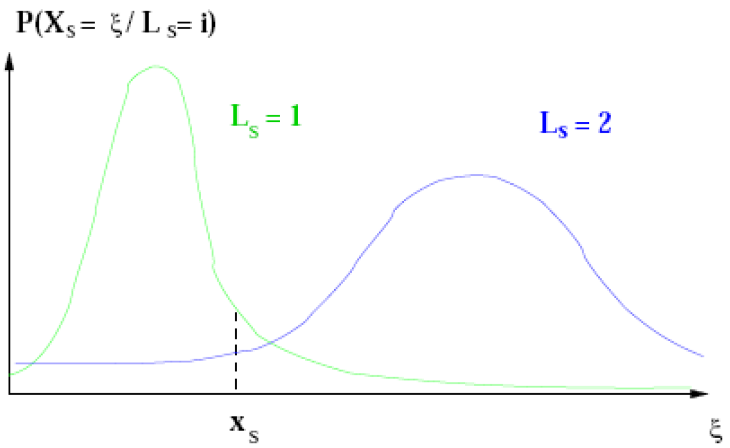
Mono-spectral case (grey-level image) (3)

How can we learn these probabilities ?

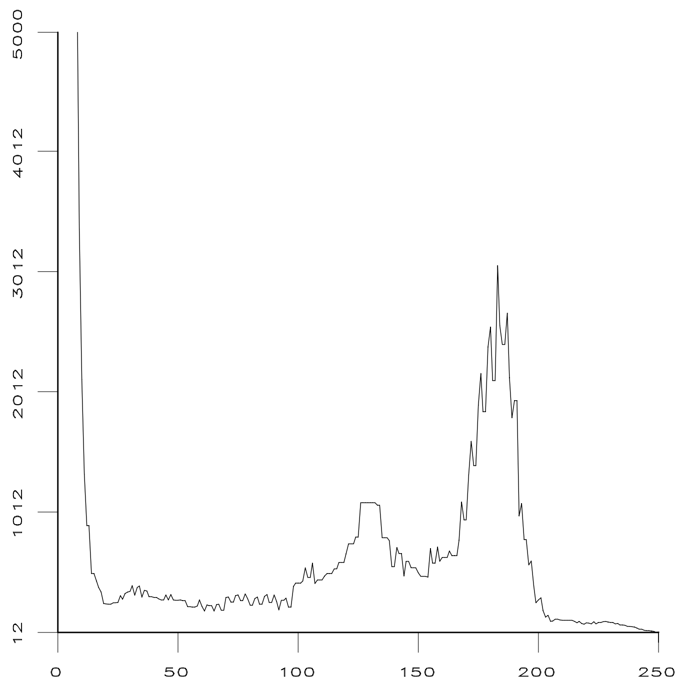
- **Learning of** $P(X_s = i)$
  - frequencies of apparition for each class
  - no knowledge : uniform distribution $(P(X_s = i) = \frac{1}{Card(\Lambda)}) \Rightarrow$ Maximum Likelihood criterion

- **Learning of** $P(Y_s = y_s | X_s = i)$

grey-level histogram for class i



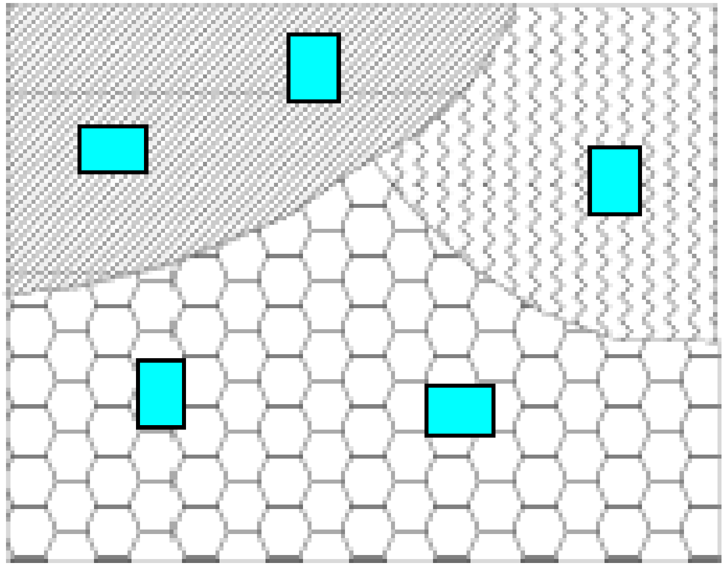
Example (brain image)



Mono-spectral case (grey-level image) (4)

◦ **Supervised learning**

- samples selection in an image
- histogram computation
- histogram filtering



◦ **Parametric case**

If there exists a parametric model for the grey-level distribution, compute the model parameters !

Ex :

- Gaussian distribution : mean, standard deviation
- Gamma distribution : mean, knowledge of the sensor parameter

Mono-spectral case (grey-level image) (5)

◦ Case of a Gaussian distribution

each class $i \in \Lambda$ is characterized by (μ_i, σ_i)

- the conditional probability is :

$$P(Y_s = y_s | X_s = i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp(-\frac{(y_s - \mu_i)^2}{2\sigma_i^2})$$

- if the classes are equiprobable:

$$P(Y_s | X_s = i) \;\; maximum \Leftrightarrow \frac{(y_s - \mu_i)^2}{2\sigma_i^2} + \ln(\sigma_i) \;\; minimum$$

- if the classes are equiprobable and have the same standard deviation (gaussian noise):

$$P(Y_s | X_s = i) \quad maximum \Leftrightarrow (y_s - \mu_i)^2 \quad minimum$$

Image classification

- **Introduction**
- **Bayesian classification**
  - Image modeling
  - Mono-spectral case
  - Multi-spectral case
  - Punctual / Contextual
- **Kmeans classification**
  - Unsupervised case
  - algorithm

Multi-spectral case

- **Vectorial observations**

$\overline{y_s} \in \{0, ..., 255\}^d$

$\downarrow$

$\overline{X_s}$

- **Bayes rule**

$$P(X_s = i | \overline{Y_s} = \overline{y_s}) = \frac{P(\overline{Y_s} = \overline{y_s} | X_s = i) P(X_s = i)}{P(\overline{Y_s} = \overline{y_s})}$$

$$x_s = argmax_{i \in \{1,...,K\}} P(\overline{Y_s} = \overline{y_s} | X_s = i) P(X_s = i)$$

$P(\overline{Y_s} = \overline{y_s} | X_s = i)$ multidimensionnal histogram for class i

Multi-spectral case

◦ Multi-variate Gaussian distribution

each class $i \in \Lambda$ is characterized by $(\overline{\mu_i}, \Sigma_i)$ (mean vector and variance-covariance matrix)

- the conditional probability is :

$$P(\overline{Y_s} = \overline{y_s} | X_s = i) = \frac{1}{\sqrt{2\pi}^d \sqrt{\det(\Sigma_i)}} \exp(-\frac{1}{2}(\overline{y_s} - \overline{\mu_i})^t \Sigma_i^{-1} (\overline{y_s} - \overline{\mu_i}))$$

- if the classes are equiprobable with the same covariance matrix:

$$P(\overline{Y_s} | X_s = i) \quad maximum \Leftrightarrow (\overline{y_s} - \overline{\mu_i})^t \Sigma^{-1} (\overline{y_s} - \overline{\mu_i}) \quad minimum$$

$\Rightarrow$ linear classifier

Image classification

◦ **Introduction**

◦ **Bayesian classification**

- Image modeling
- Mono-spectral case
- Multi-spectral case
- Punctual / Contextual

◦ **Kmeans classification**

- Unsupervised case
- algorithm

Contextual / punctual classification

◦ **Global classification**

$y = \{y_s\}_{s \in S}$ (observed image), $Y = \{Y_s\}_{s \in S}$ (random field)

$x = \{x_s\}_{s \in S}$ (searched classification), $X = \{X_s\}_{s \in S}$ (random field)

$$P(X = x | Y = y) = \frac{P(Y = y | X = x) P(X = x)}{P(Y = y)}$$

◦ **Independance assumption for** $P(Y = y | X = x)$

$$P(Y = y | X = x) = \prod_{s \in S} P(Y_s = y_s | X_s = x_s)$$

◦ **Independance assumption for** $P(X = x)$

$$P(X = x) = \prod_{s \in S} P(X_s = x_s)$$

$\Rightarrow P(X = x | Y = y) \propto \prod_{s \in S} P(Y_s = y_s | X_s) P(X_s = x_s) \Rightarrow$ punctual classif.!

Contextual / punctual classification

◦ **Prior knowledge on $P(X)$**

- independence assumption not verified in practice: images are smooth with strong spatial coherency (image description = smooth areas)
- BUT the coherency is at a local scale ⇒ introduction of contextual knowledge

Markov random fields ⇒ smoothness of the solution, local spatial coherency in the result !

Image classification

◦ **Introduction**

◦ **Bayesian classification**

- Image modeling
- Mono-spectral case
- Multi-spectral case
- Punctual / Contextual

◦ **Kmeans classification**

- Unsupervised case
- algorithm

Unsupervised case

◦ **Bayesian case with gaussian distribution and same std**

$$P(Y_s|X_s = i) \quad maximum \Leftrightarrow ||y_s - \mu_i||^2 \quad minimum$$

◦ **Unsupervised classification**

μ_i unknown $\Rightarrow$

- classify the pixels using some initial μ_i^0
- compute new μ_i^1 with the empirical mean of the classified pixels
- iterate until no modification in μ_i^k

$\Rightarrow$kind of bayesian classification with changing means

Unsupervised case

◦ **K-means algorithm**

- choose $\mu_1^0, \mu_2^0, ..., \mu_K^0$

At iteration k:

- $\forall s \in S \quad l_s = argmin_{i \in \Lambda} ||y_s - \mu_i^k||^2$
- $\forall i \in \{1, ..., K\} \quad \mu_i^{k+1} = \frac{1}{card(R_i)} \sum_{s, l_s = i} x_s$
- if $\mu_i^k \neq \mu_i^{k+1}$ iterate

◦ **Drawbacks**

- no proof of convergence to the optimal solution
- influence of the initial means

Unsupervised case

◦ **K-means algorithm for grey-level images**

in 1D, everything can be done with the histogram!

f_n frequency of grey-level n in the image

Ex for 2 classes:

initial values of the centers $m_1^0 = 10$; $m_2^0 = 120$

initial classification: thresholding with $t = 65$ (histogram)

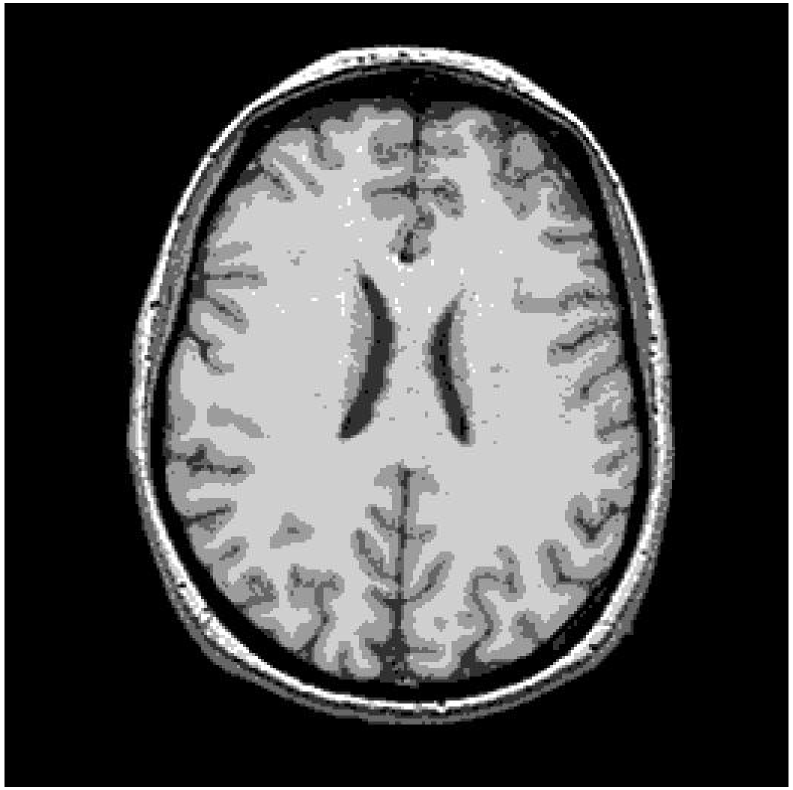
computation of new means :

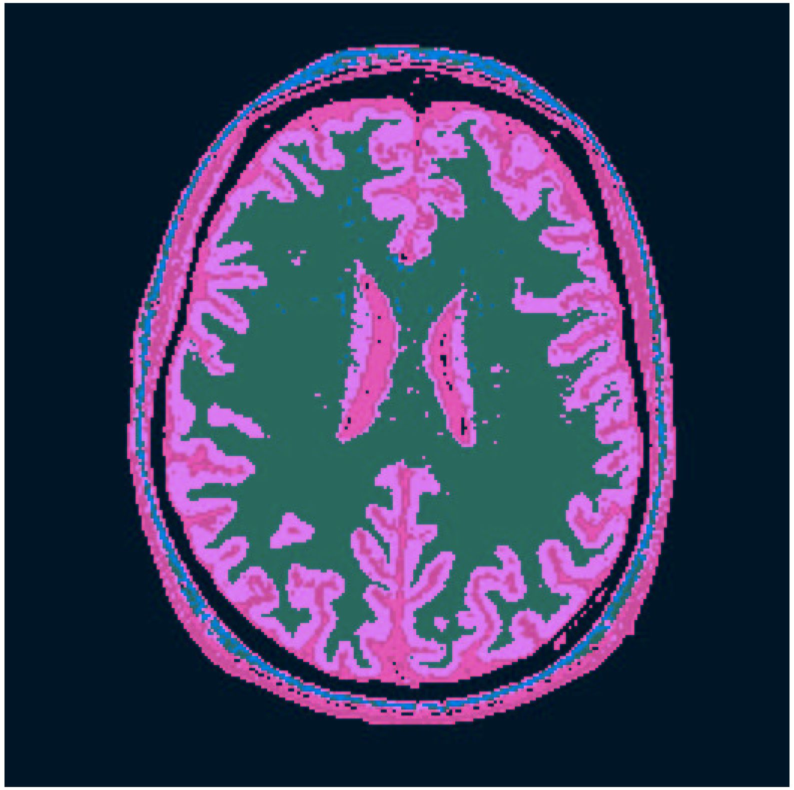
$$\mu_1^1 = \frac{1}{\sum_{n=0}^{n=t} f_n} \sum_{n=0}^{n=t} n f_n$$

$$\mu_2^1 = \frac{1}{\sum_{n=t}^{n=N} f_n} \sum_{n=t}^{n=N} n f_n$$

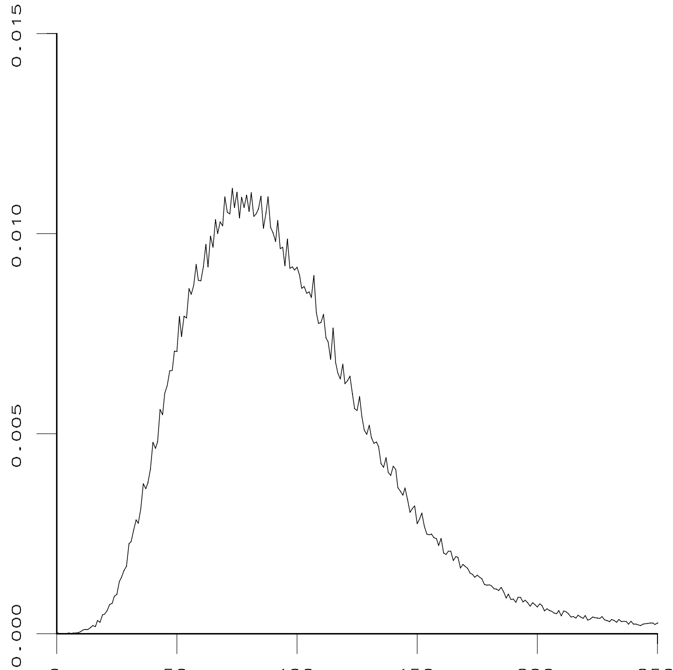
- multi-channels : in nD, high cost of storage ⇒ updating of a classified image
- Application to high dimensionnal space : reduction with ACP

Brain image

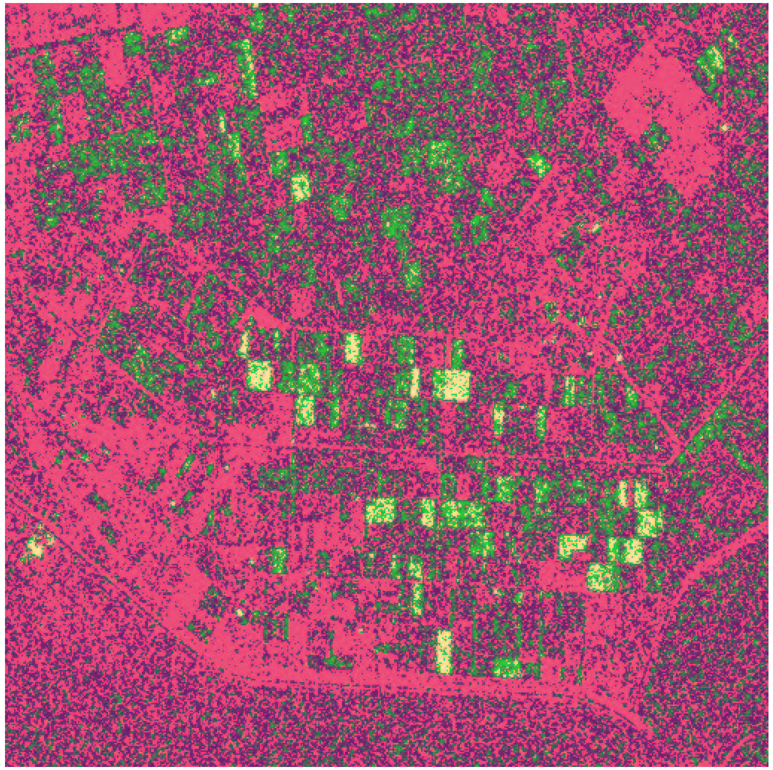


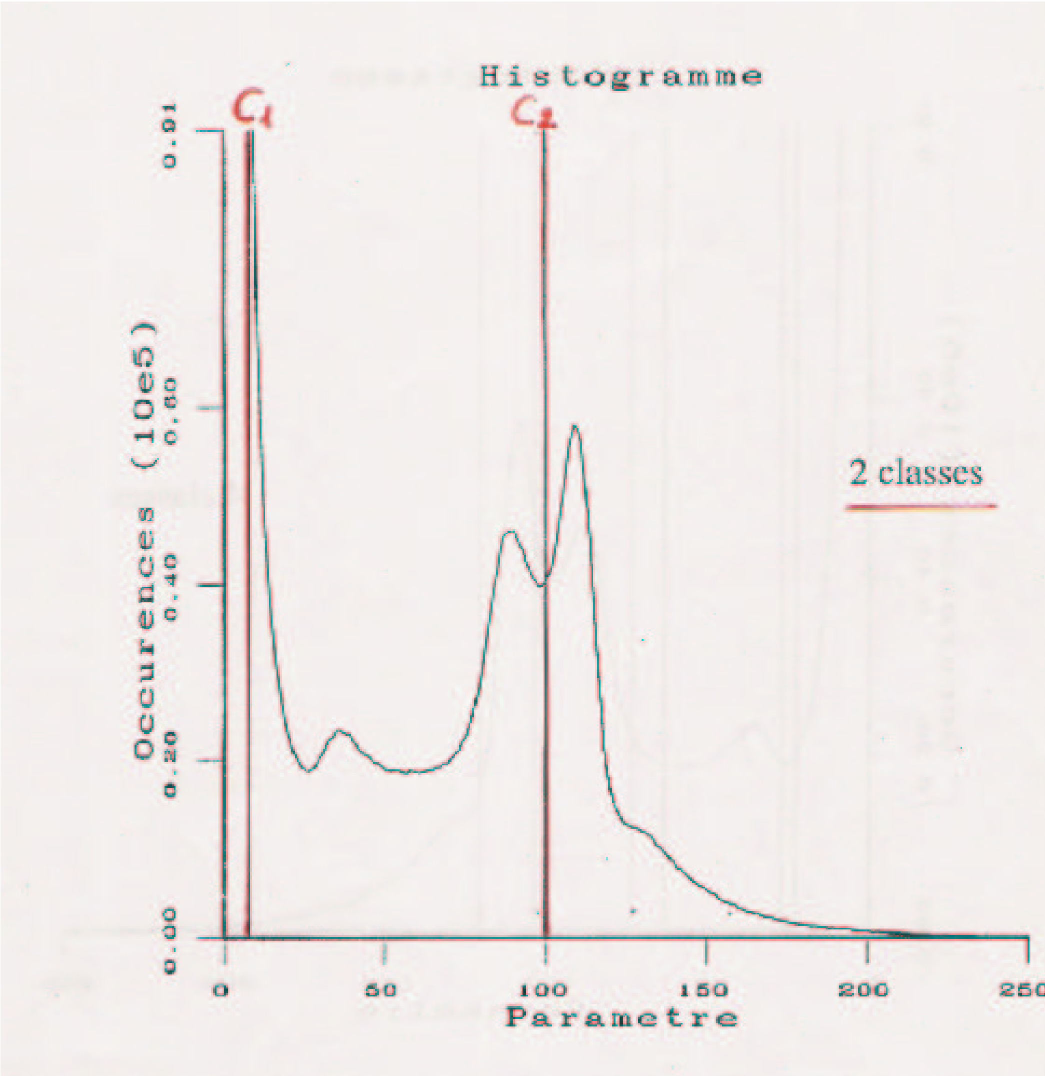


SAR image



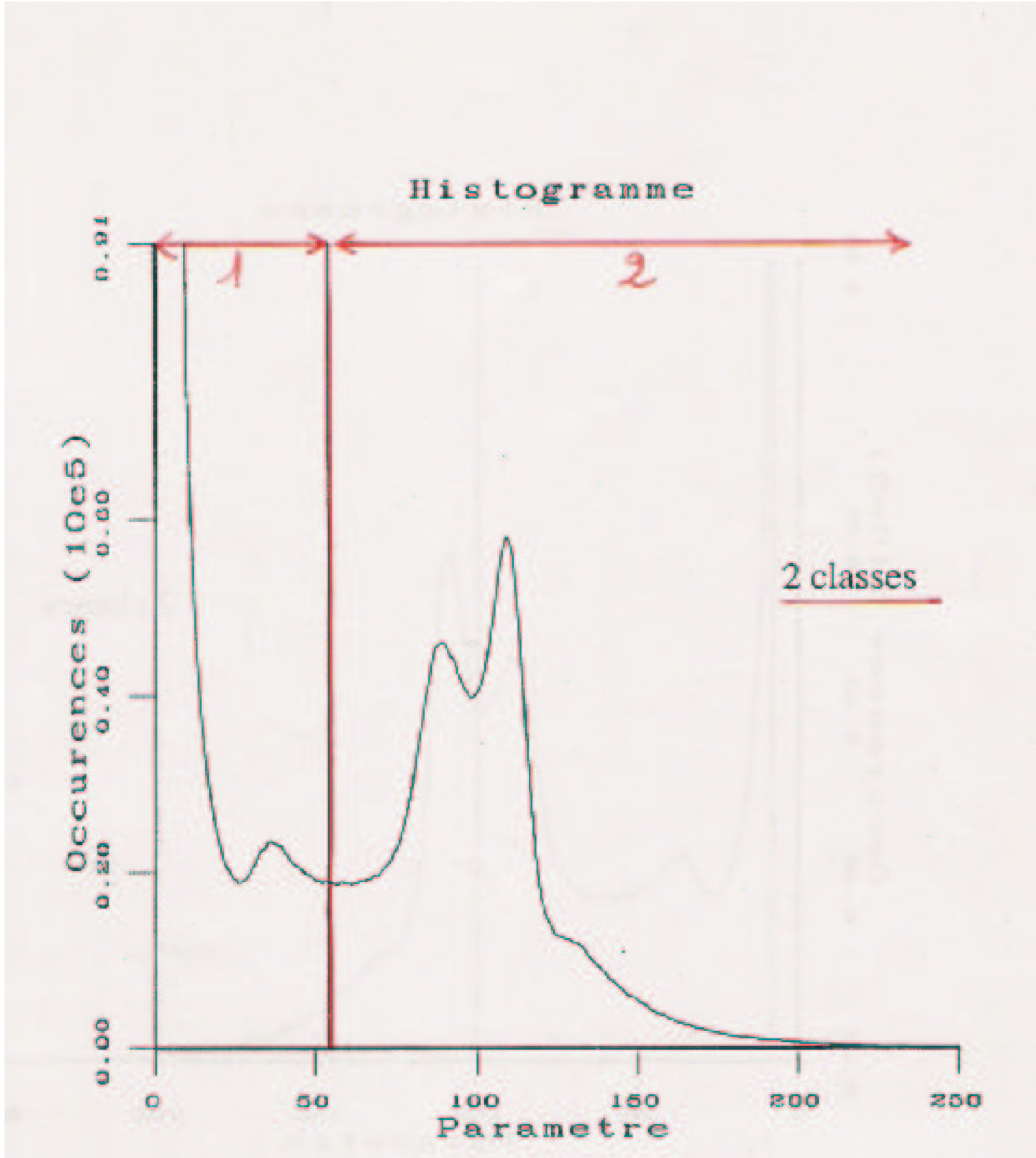
Imagerie radar



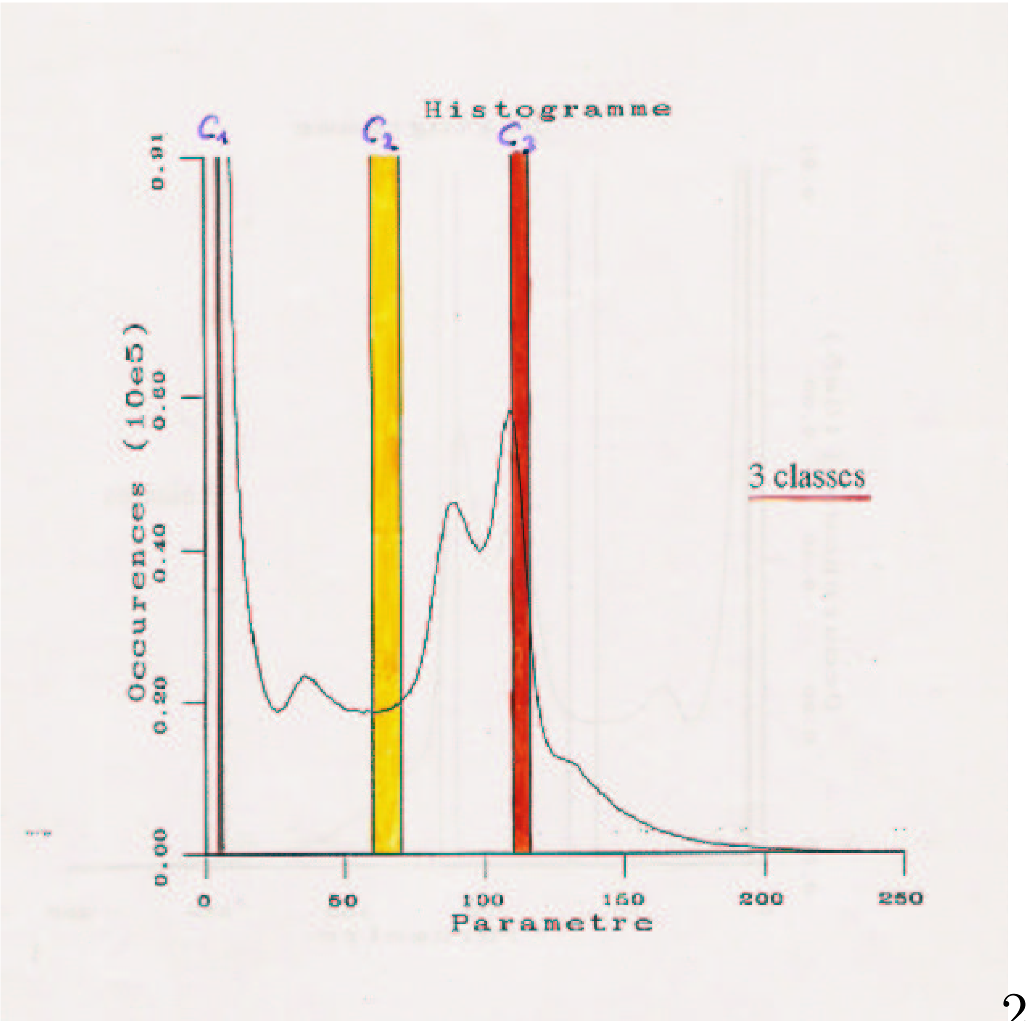


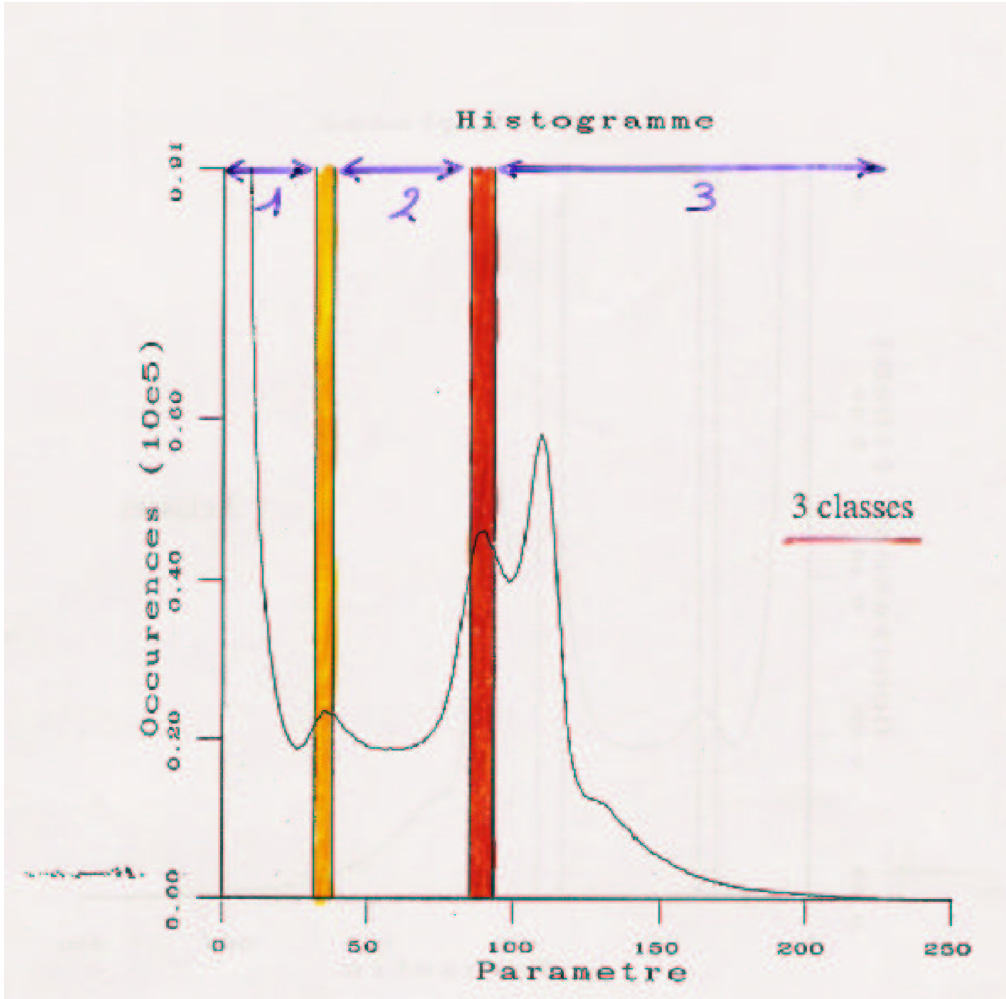
Histogramme
C1
C2
Occurences (10e5)
0.00
0.20
0.40
0.60
0.91
0
50
100
150
200
250
Parametre
2 classes

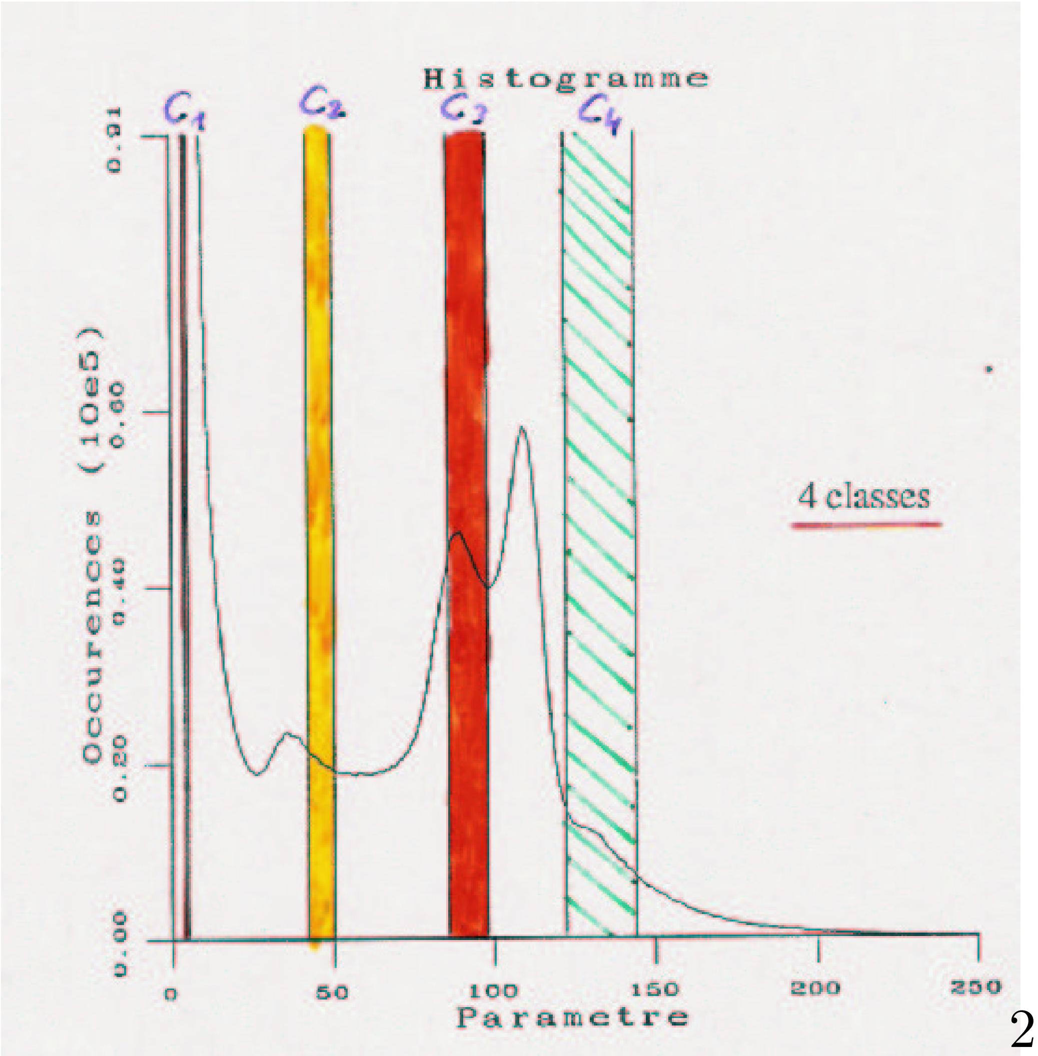
27-1

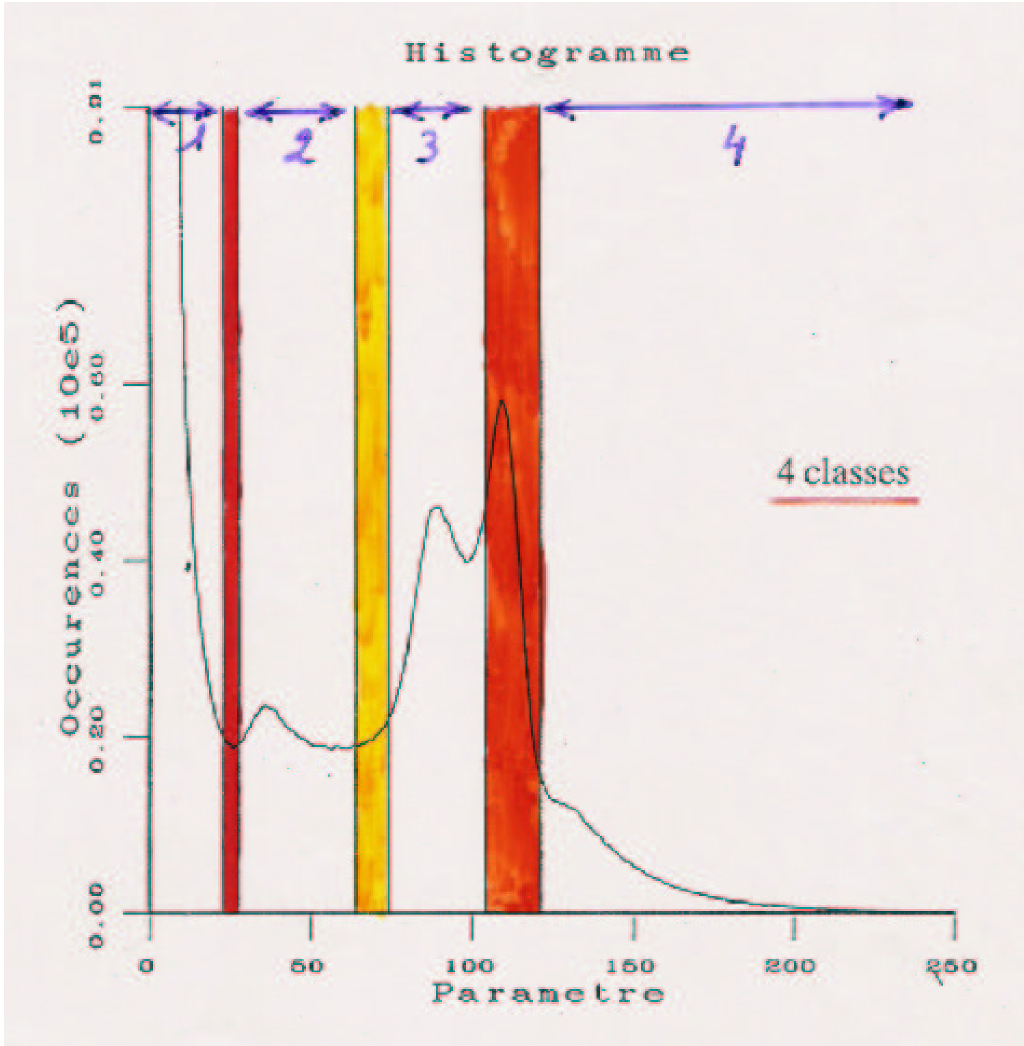


Histogramme
1
2
Occurences (10e5)
0.00
0.20
0.40
0.60
0.91
0
50
100
150
200
250
Parametre
2 classes


Histogramme
C1
C2
C3
Occurences (10e5)
0.00
0.20
0.40
0.60
0.91
0
50
100
150
200
250
Parametre
3 classes


Histogramme
1
2
3
Occurences (10e5)
0.00
0.20
0.40
0.60
0.91
0
50
100
150
200
250
Parametre
3 classes


Histogramme
C1
C2
C3
C4
Occurences (10e5)
0.00
0.20
0.40
0.60
0.81
4 classes
0
50
100
150
200
250
Parametre


Histogramme
1
2
3
4
Occurences (10e5)
0.00
0.20
0.40
0.60
0.81
4 classes
0
50
100
150
200
250
Parametre