



# Concerto pour violons en stats majeures



E. Le Penec (SELECT - INRIA Saclay / Université Paris Sud)

et

S. Cohen (IPANEMA - Soleil Saclay)

avec

G. Celeux et P. Massart (SELECT - INRIA Saclay / Université Paris Sud)

et C. Maugis (INSA Toulouse)



Unithé ou café? / 01 octobre 2010

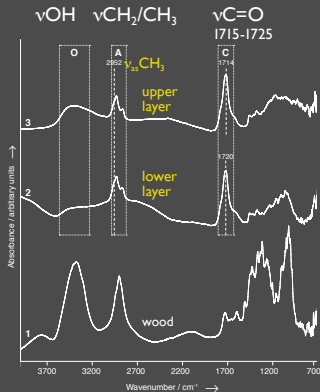


# A. Stradivari (1644 - 1737)

Provigny (1716)



A. Giordan © Cité de la Musique



SOLEIL  
SYNCHROTRON

4 / 8 cm<sup>-1</sup> resolution  
64 / 128 scans  
typ. 1 min/sp, 400sp

very simple process  
no protein (amide I, amide II)  
no gums, nor waxes  
@SOLEIL: SMIS



J.-P. Echard, L. Bertrand, A. von Bohlen, A.-S. Le Hô, C. Paris, L. Bellot-Gurlet, B. Soulier, A. Lattuari-Derieux, S. Thao, L. Robinet, B. Lavédrine, and S. Vaiedelich. *Angew. Chem. Int. Ed.*, 49(1), 197-201, 2010.



# Segmentation d'images hyperspectrales

# Segmentation d'images hyperspectrales

- Données :
  - image de taille  $n$  comprise entre  $\sim 1000$  et  $\sim 100000$  pixels,
  - spectres  $\mathcal{S}$  de  $\sim 1024$  points,
  - résolution  $\sim 4/8 \text{ cm}^{-1}$  (10 fois meilleure dans le visible),
  - possibilité de mesurer de très nombreux spectres par minute...

# Segmentation d'images hyperspectrales

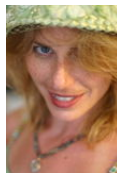
- Données :
  - image de taille  $n$  comprise entre  $\sim 1000$  et  $\sim 100000$  pixels,
  - spectres  $\mathcal{S}$  de  $\sim 1024$  points,
  - résolution  $\sim 4/8 \text{ cm}^{-1}$  (10 fois meilleure dans le visible),
  - possibilité de mesurer de très nombreux spectres par minute...
- Objectifs immédiats :
  - segmentation automatique de ces images,
  - sans intervention humaine,
  - aide à l'analyse des résultats.

# Segmentation d'images hyperspectrales

- Données :
  - image de taille  $n$  comprise entre  $\sim 1000$  et  $\sim 100000$  pixels,
  - spectres  $\mathcal{S}$  de  $\sim 1024$  points,
  - résolution  $\sim 4/8 \text{ cm}^{-1}$  (10 fois meilleure dans le visible),
  - possibilité de mesurer de très nombreux spectres par minute...
- Objectifs immédiats :
  - segmentation automatique de ces images,
  - sans intervention humaine,
  - aide à l'analyse des résultats.
- Objectifs lointains :
  - classification automatique,
  - interprétation...

# Classification “non supervisée”

# Classification “non supervisée”



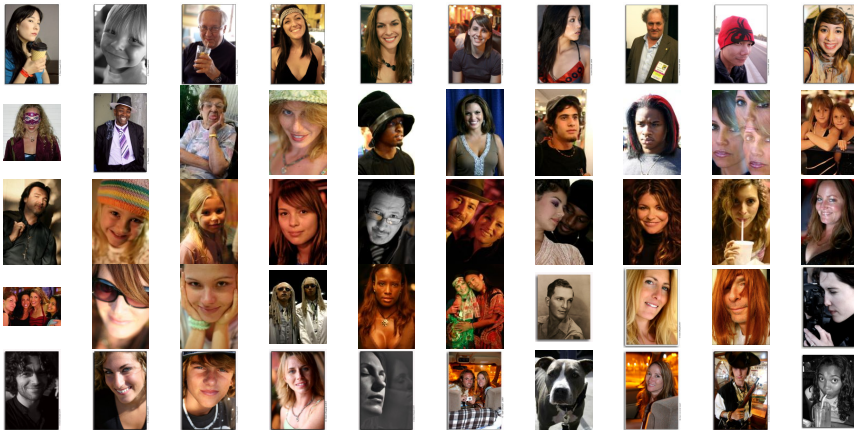


# Classification “non supervisée”



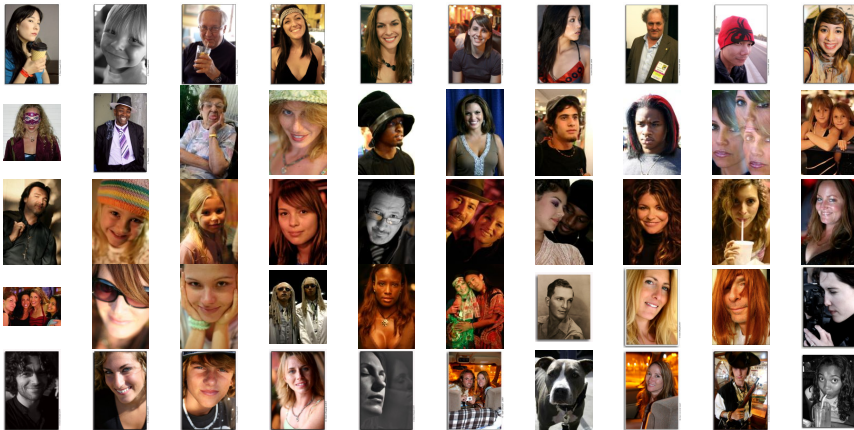
- Classer des objets sans aucune autre indication.

# Classification “non supervisée”



- Classifier des objets sans aucune autre indication.

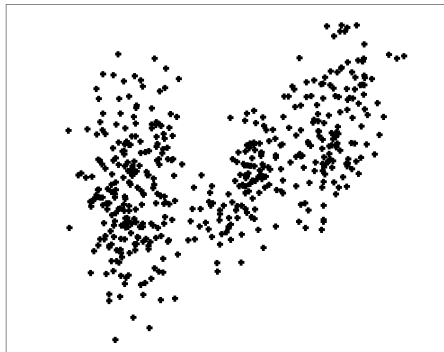
# Classification “non supervisée”



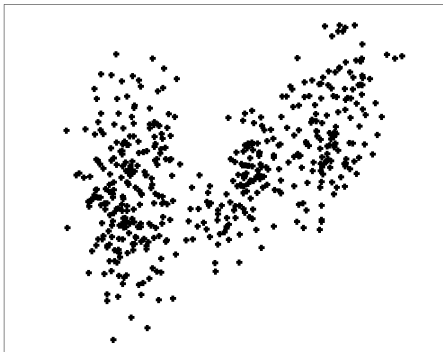
- Classer des objets sans aucune autre indication.
- Mal posé : pas une “bonne réponse” unique !

# Un problème “jouet”

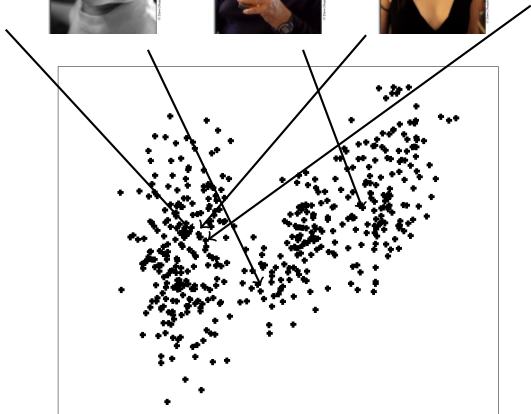
# Un problème “jouet”



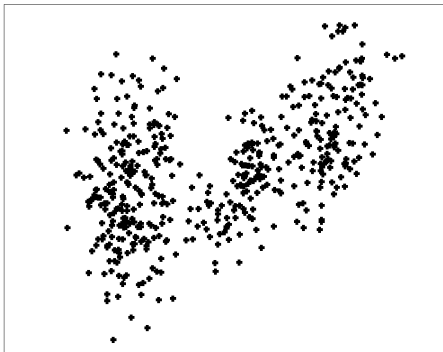
# Un problème “jouet”



# Un problème “jouet”

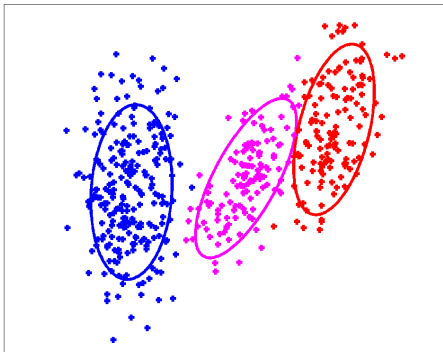


# Un problème “jouet”

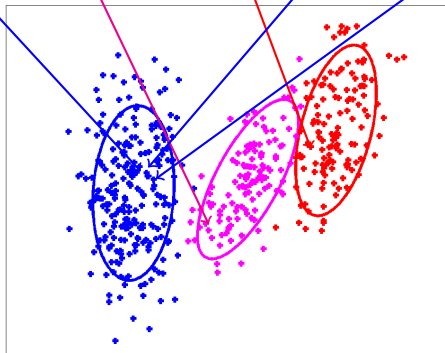




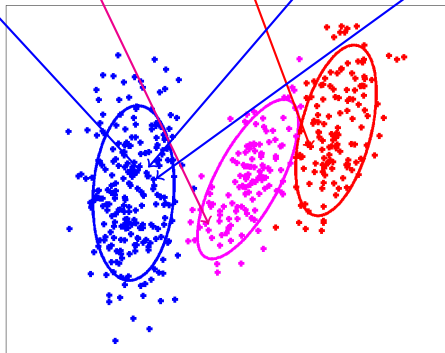
# Un problème “jouet”



# Un problème “jouet”



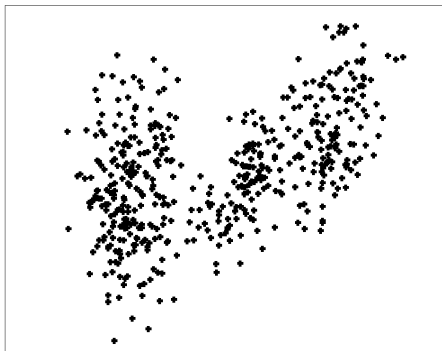
# Un problème “jouet”



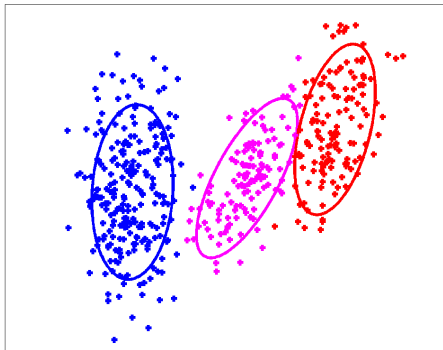
- Représentation : utilisation d'une correspondance entre les objets et des points dans un espace de grande dimension.

# Un cadre "aléatoire"

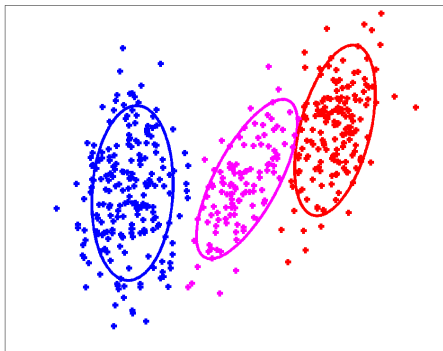
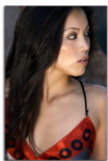
# Un cadre "aléatoire"



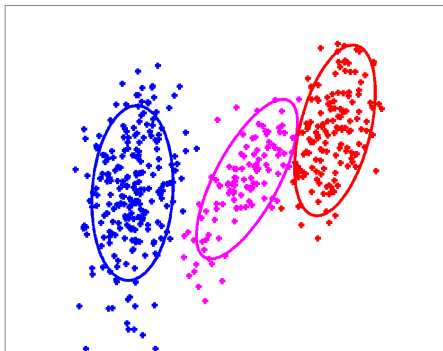
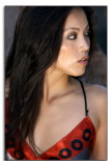
# Un cadre "aléatoire"



# Un cadre "aléatoire"

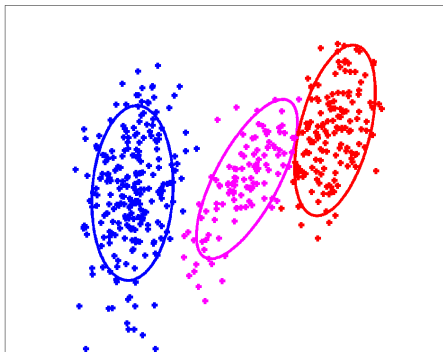
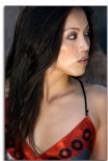


# Un cadre "aléatoire"





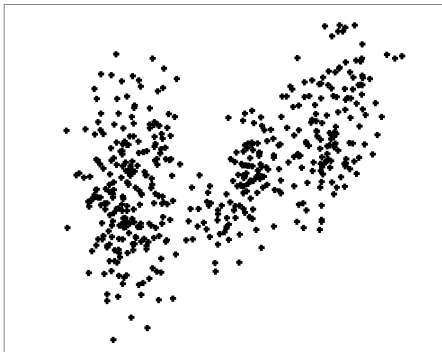
# Un cadre "aléatoire"



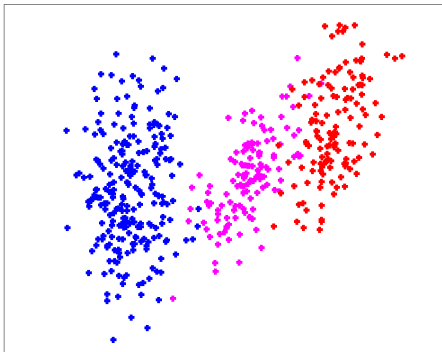
- Robustesse vis à vis des données.
- Modélisation : échantillons  $X_1, \dots, X_n$  i.i.d. de loi de densité  $s$ .

# Modélisation “stochastique”

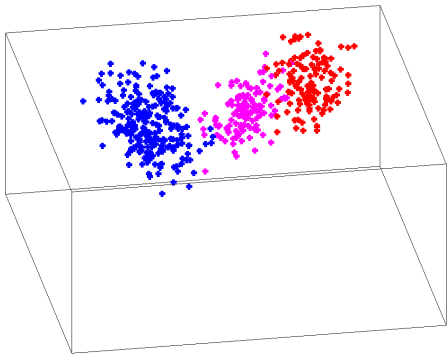
# Modélisation “stochastique”



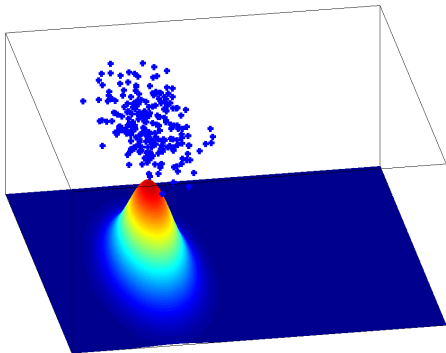
# Modélisation “stochastique”



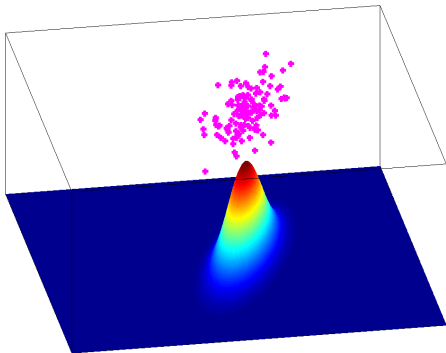
# Modélisation “stochastique”



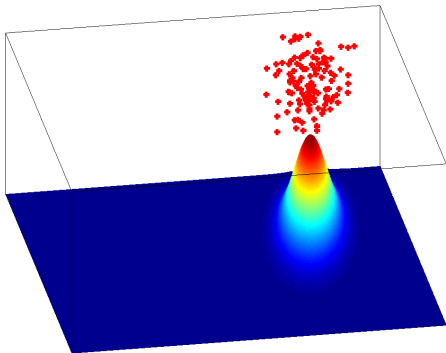
# Modélisation “stochastique”



# Modélisation “stochastique”

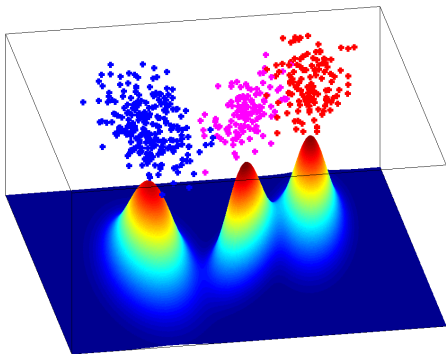


# Modélisation “stochastique”

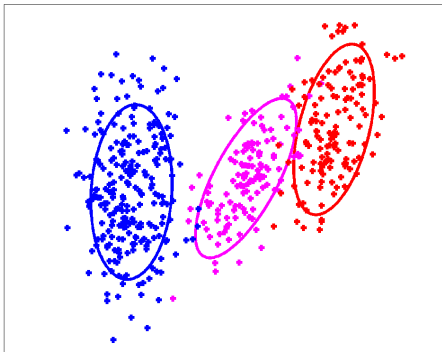




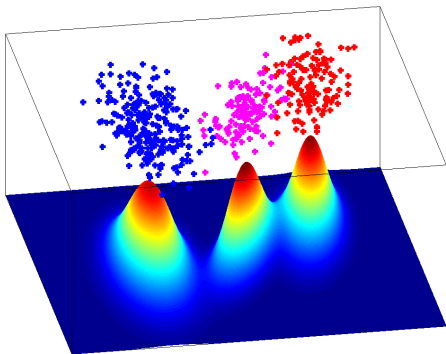
# Modélisation “stochastique”



# Modélisation “stochastique”



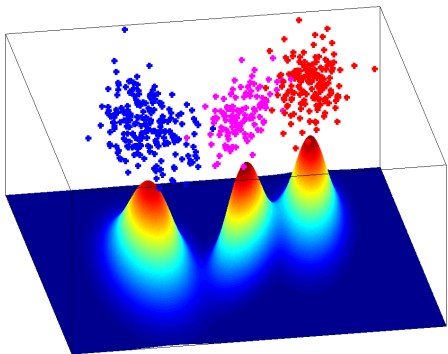
# Modélisation “stochastique”



- Modèle : mélange de gaussiennes à  $K$  classes.
- Densité du mélange :

$$s_{K,(\pi_k, \mu_k, \Sigma_k)}(x) = \sum_{k=1}^K \pi_k \frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} e^{-\frac{1}{2}(x-\mu_k)^t \Sigma_k^{-1} (x-\mu_k)} dx$$

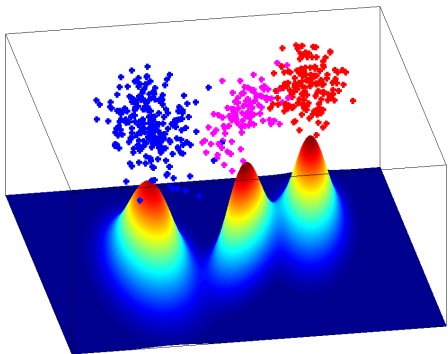
# Modélisation “stochastique”



- Modèle : mélange de gaussiennes à  $K$  classes.
- Densité du mélange :

$$s_{K,(\pi_k, \mu_k, \Sigma_k)}(x) = \sum_{k=1}^K \pi_k \frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} e^{-\frac{1}{2}(x-\mu_k)^t \Sigma_k^{-1} (x-\mu_k)} dx$$

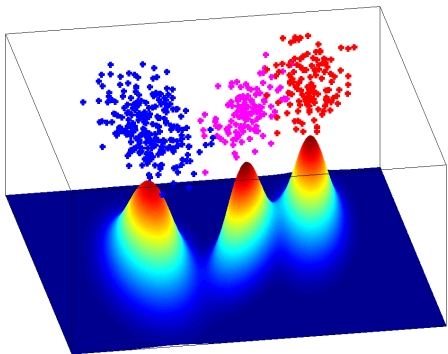
# Modélisation “stochastique”



- Modèle : mélange de gaussiennes à  $K$  classes.
- Densité du mélange :

$$s_{K,(\pi_k, \mu_k, \Sigma_k)}(x) = \sum_{k=1}^K \pi_k \frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} e^{-\frac{1}{2}(x-\mu_k)^t \Sigma_k^{-1} (x-\mu_k)} dx$$

# Modélisation “stochastique”

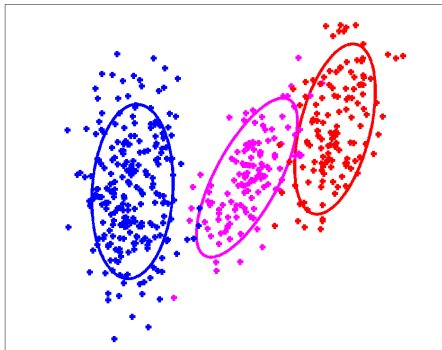


- Modèle : mélange de gaussiennes à  $K$  classes.
- Densité du mélange :

$$s_{K,(\pi_k, \mu_k, \Sigma_k)}(x) = \sum_{k=1}^K \pi_k \frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} e^{-\frac{1}{2}(x-\mu_k)^t \Sigma_k^{-1} (x-\mu_k)} dx$$

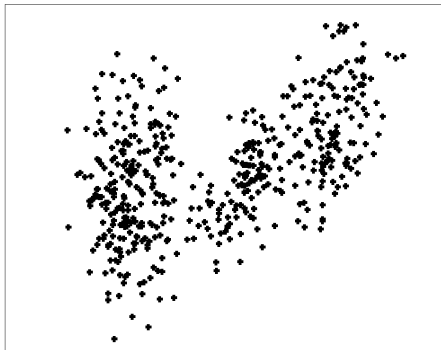
# Estimation “statistique”

# Estimation “statistique”

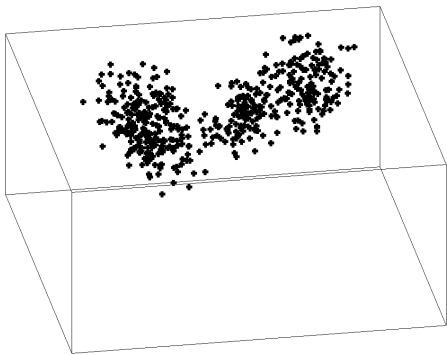




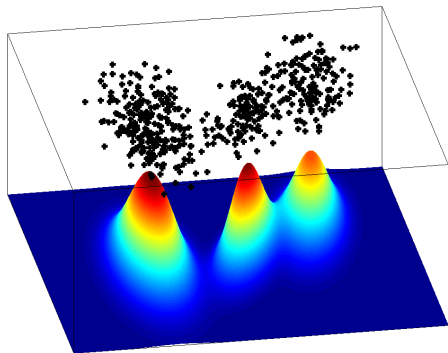
# Estimation “statistique”



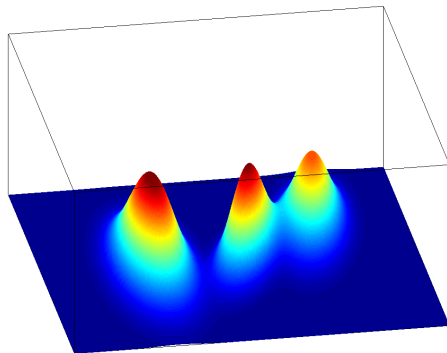
# Estimation “statistique”



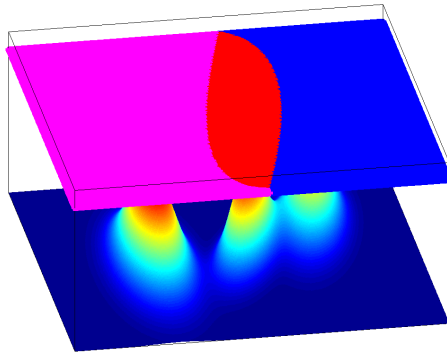
# Estimation “statistique”



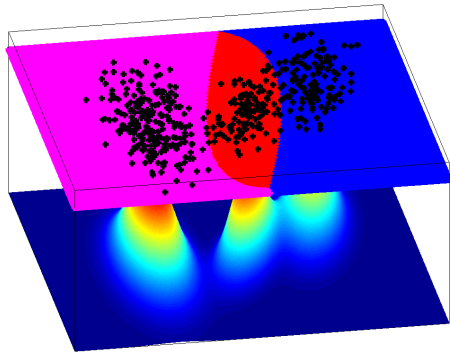
# Estimation “statistique”



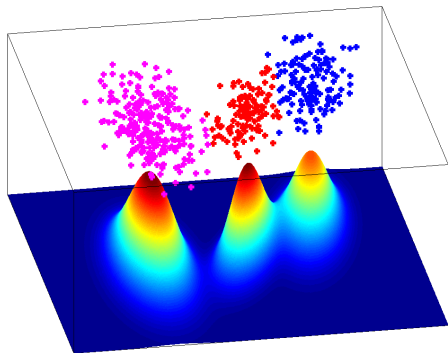
# Estimation “statistique”



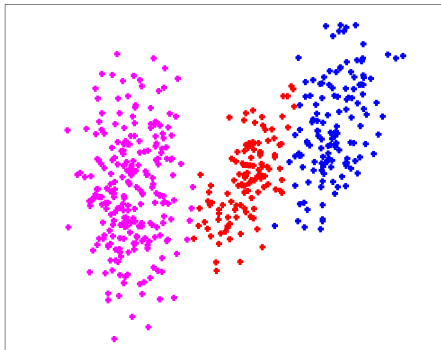
# Estimation “statistique”



# Estimation “statistique”

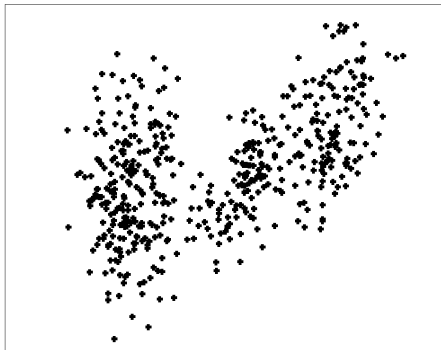


# Estimation “statistique”

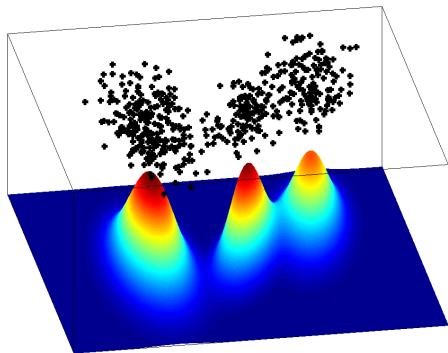




# Estimation “statistique”



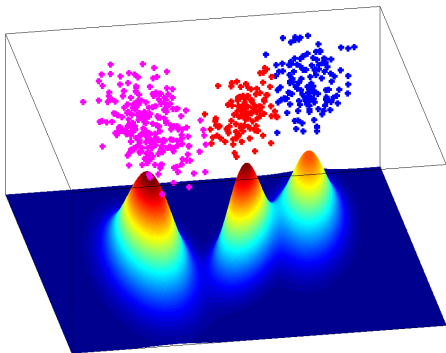
# Estimation “statistique”



- Estimation des  $\pi_k$ ,  $\hat{\mu}_k$  et  $\hat{\Sigma}_k$  par maximum de vraisemblance :

$$(\hat{\pi}_k, \hat{\mu}_k, \hat{\Sigma}_k) = \operatorname{argmax} \sum_{i=1}^n \log s_{K, (\pi_k, \mu_k, \Sigma_k)}(X_i)$$

# Estimation “statistique”



- Estimation des  $\pi_k$ ,  $\hat{\mu}_k$  et  $\hat{\Sigma}_k$  par maximum de vraisemblance :

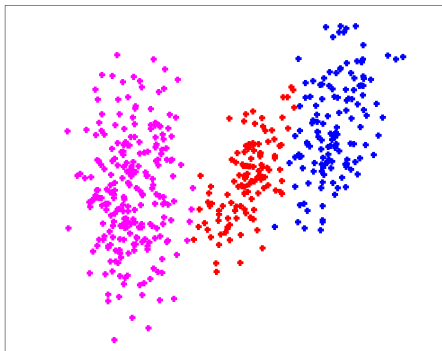
$$(\hat{\pi}_k, \hat{\mu}_k, \hat{\Sigma}_k) = \operatorname{argmax} \sum_{i=1}^n \log s_{K, (\pi_k, \mu_k, \Sigma_k)}(X_i)$$

- Estimation de  $\hat{k}(x)$  par maximum à posteriori :

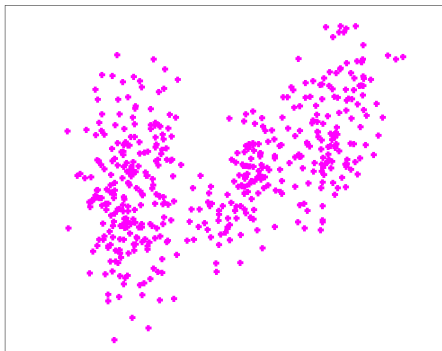
$$\hat{k}(x) = \operatorname{argmax} \hat{\pi}_k \frac{1}{\sqrt{(2\pi)^d |\hat{\Sigma}_k|}} e^{-\frac{1}{2}(x-\hat{\mu}_k)^t \hat{\Sigma}_k^{-1} (x-\hat{\mu}_k)}$$

Combien de classes ?

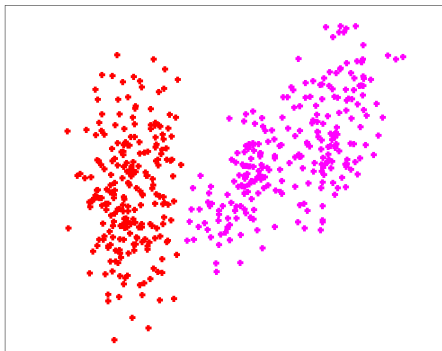
Combien de classes ?



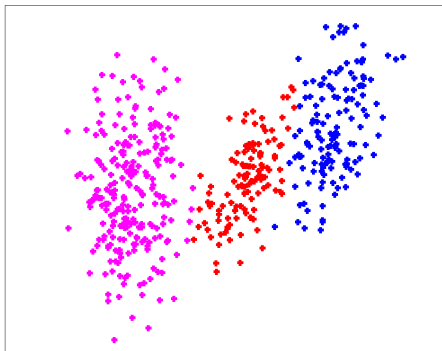
# Combien de classes ?



Combien de classes ?

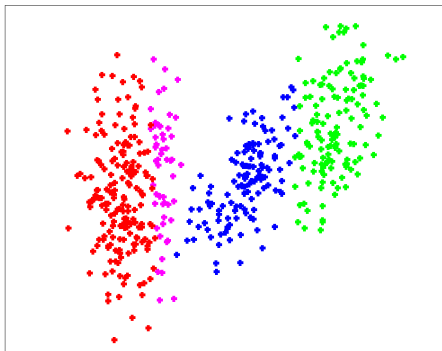


# Combien de classes ?

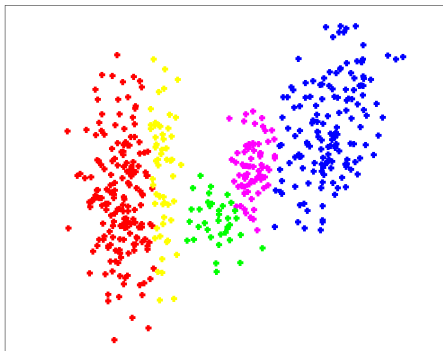




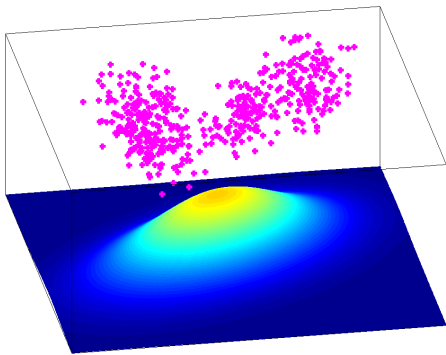
# Combien de classes ?



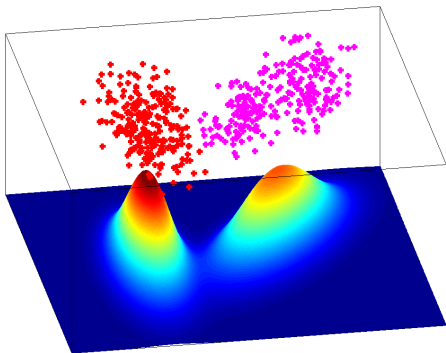
Combien de classes ?



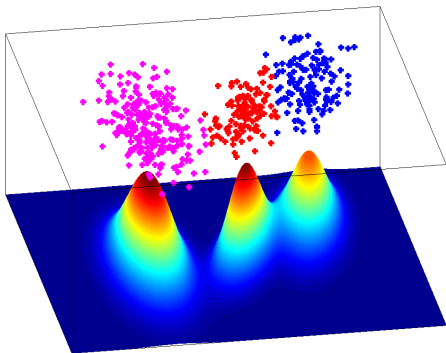
Combien de classes ?



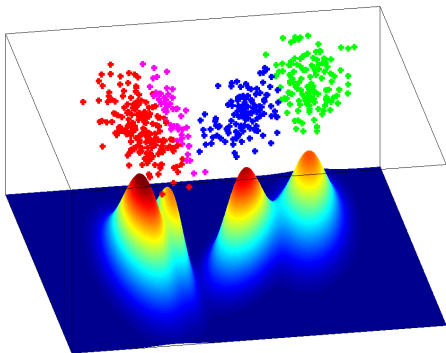
Combien de classes ?



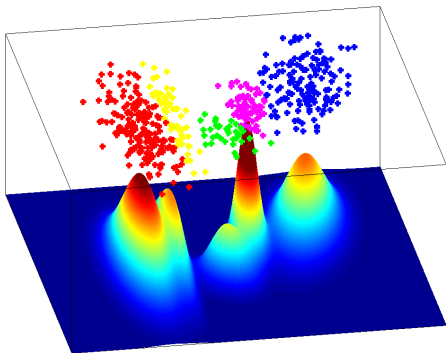
Combien de classes ?



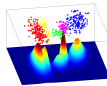
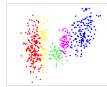
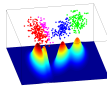
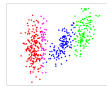
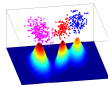
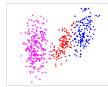
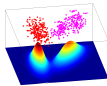
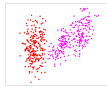
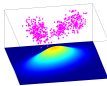
Combien de classes ?



Combien de classes ?

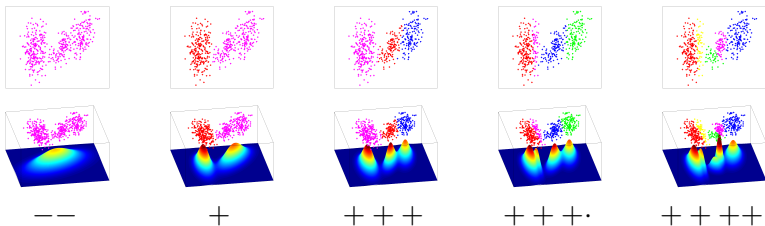


# Combien de classes ?

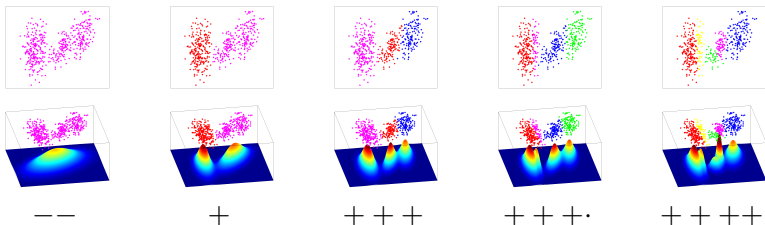




# Combien de classes ?

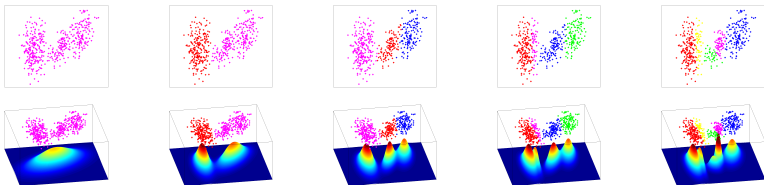


# Combien de classes ?



- Question difficile où la vraisemblance (la fidélité) ne suffit pas !

# Combien de classes ?



Fidélité

--

+

+++

+++.

++++

Simplicité

++++

+++

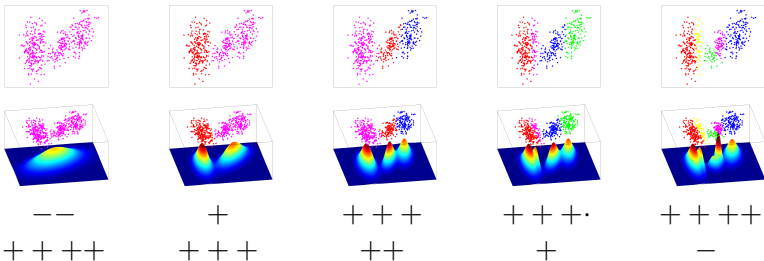
++

+

-

● Question difficile où la vraisemblance (la fidélité) ne suffit pas !

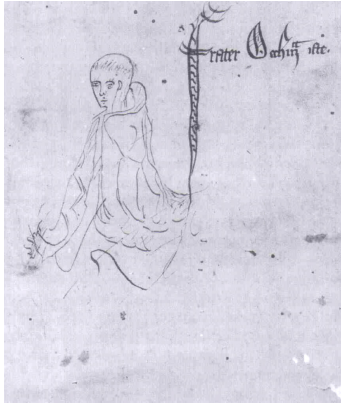
# Combien de classes ?



- Question difficile où la vraisemblance (la fidélité) ne suffit pas !
- Prise en compte de la complexité du modèle ?

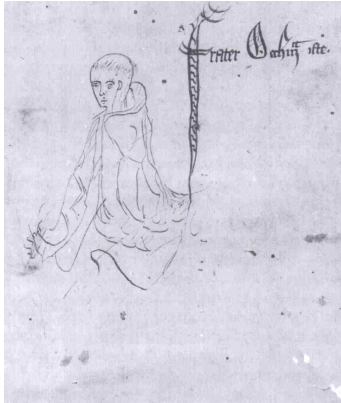
# Le rasoir d'Ockham

# Le rasoir d'Ockham



*Les multiples ne doivent pas être utilisés sans nécessité.*  
Guillaume d'Ockham (~ 1285 - 1347)

# Le rasoir d'Ockham



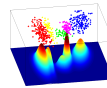
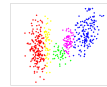
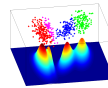
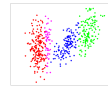
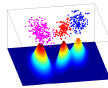
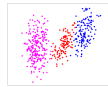
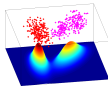
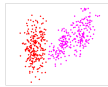
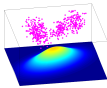
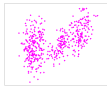
*Les multiples ne doivent pas être utilisés sans nécessité.*  
Guillaume d'Ockham (~ 1285 - 1347)

- Rasoir d'Ockham (principe de simplicité) : il ne faut pas ajouter des hypothèses, si celles utilisées suffisent déjà !
- Compromis entre pouvoir d'explication et simplicité.

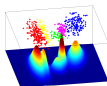
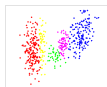
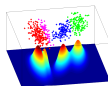
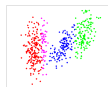
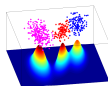
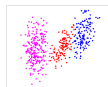
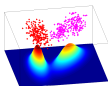
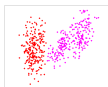
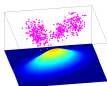
# Sélection par pénalisation



# Sélection par pénalisation



# Sélection par pénalisation



Vraisemblance

--

+

+++

+++.

++++

Simplicité

++++

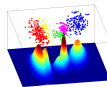
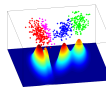
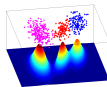
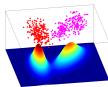
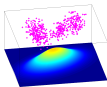
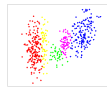
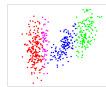
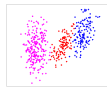
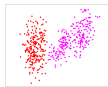
+++

++

+

-

# Sélection par pénalisation



Vraisemblance

--

+

+++

+++.

++++

+ Simplicité

++++

+++

++

+

-

= Compromis

++

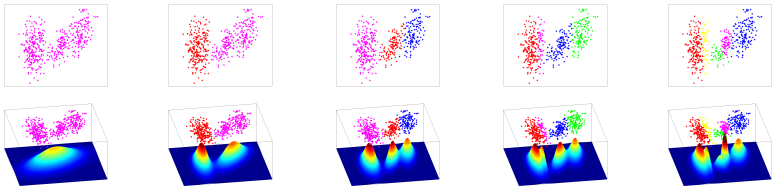
++++

+++++

++++.

+++

# Sélection par pénalisation



Vraisemblance

--

+

+++

+++.

++++

+ Simplicité

++++

+++

++

+

-

= Compromis

++

++++

+++++

++++.

+++

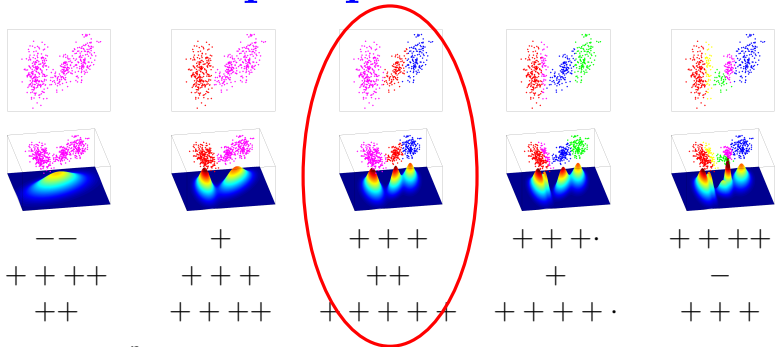
● Vraisemblance :  $\sum_{i=1}^n \log \hat{s}_K(X_i)$ .

● Simplicité :  $-\lambda \text{Dim}(S_K)$  (beaucoup de théorie derrière).

● Estimateur pénalisé :

$$\text{argmax} \underbrace{\sum_{i=1}^n \log \hat{s}_K(X_i)}_{\text{Vraisemblance}} - \underbrace{\lambda \text{Dim}(S_K)}_{\text{Pénalité}}$$

# Sélection par pénalisation



● Vraisemblance :  $\sum_{i=1}^n \log \hat{s}_K(X_i)$ .

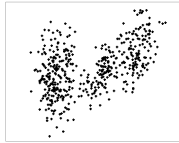
● Simplicité :  $-\lambda \text{Dim}(S_K)$  (beaucoup de théorie derrière).

● Estimateur pénalisé :

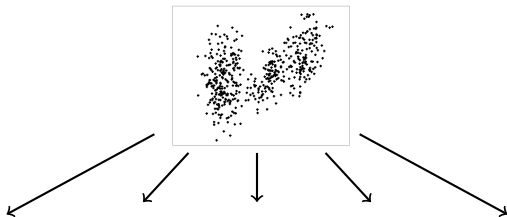
$$\text{argmax} \underbrace{\sum_{i=1}^n \log \hat{s}_K(X_i)}_{\text{Vraisemblance}} - \underbrace{\lambda \text{Dim}(S_K)}_{\text{Pénalité}}$$

# Méthodologie

# Méthodologie

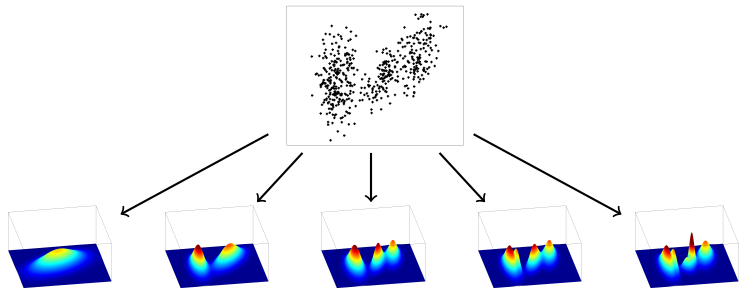


# Méthodologie

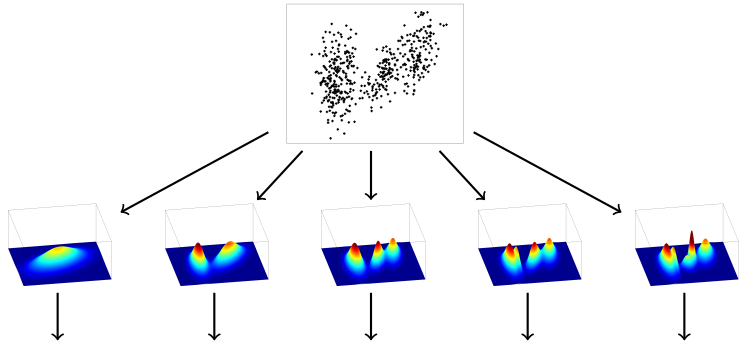




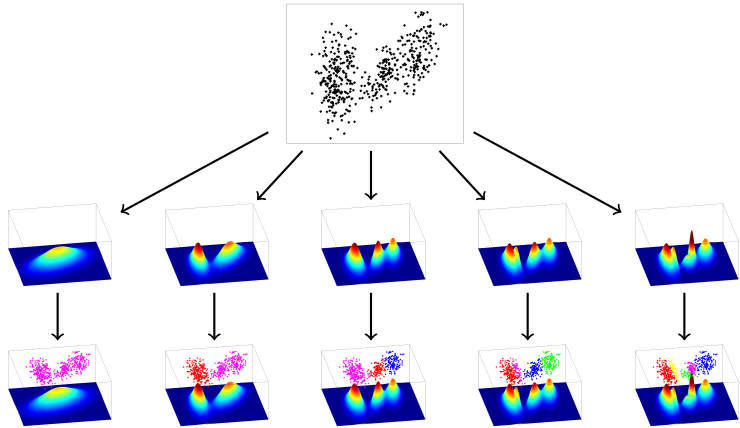
# Méthodologie



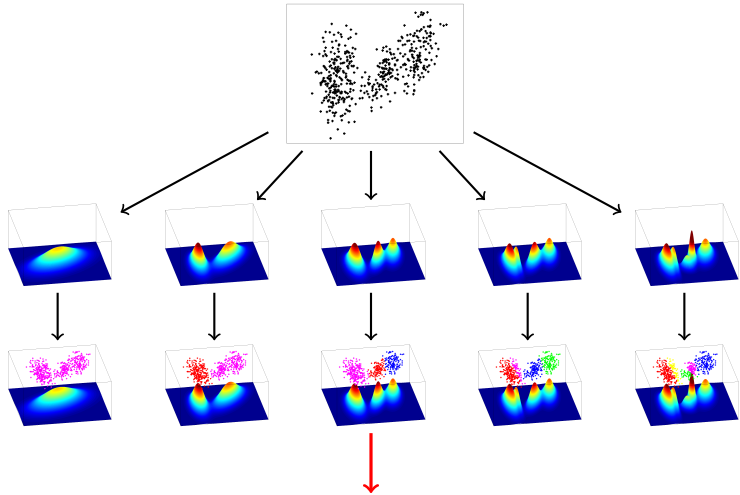
# Méthodologie



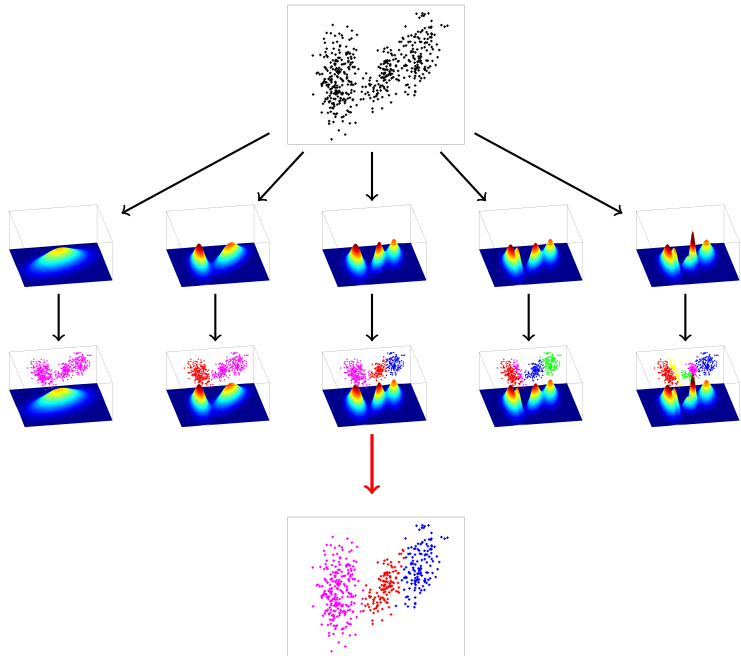
# Méthodologie



# Méthodologie

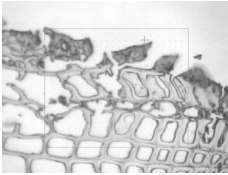


# Méthodologie

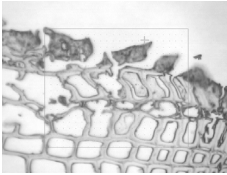


Retour à nos violons

# Retour à nos violons



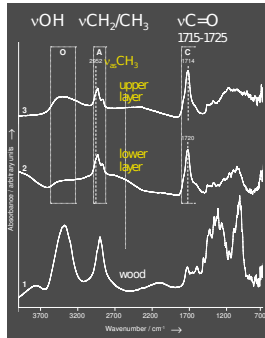
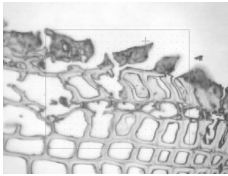
# Retour à nos violons



segmentation

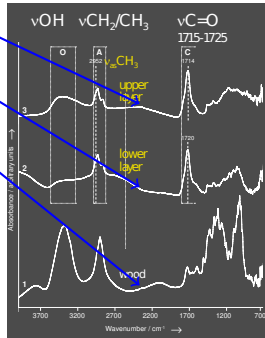
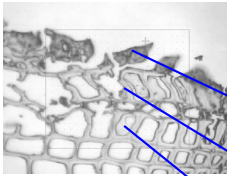


# Retour à nos violons



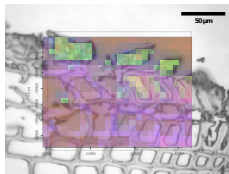
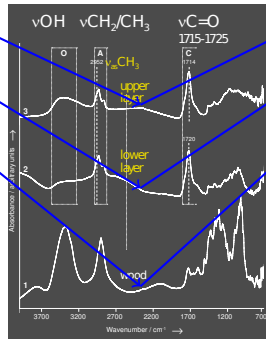
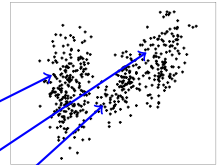
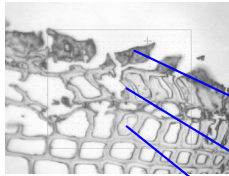
gmentation

# Retour à nos violons

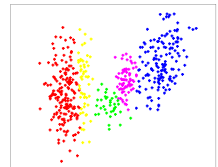
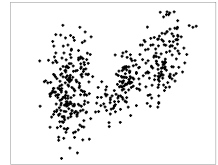
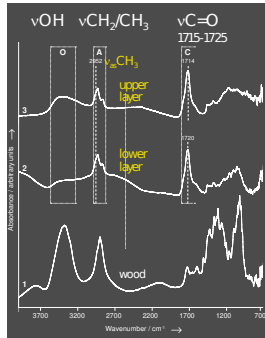
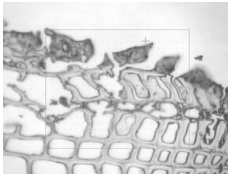


gmentation

# Retour à nos violons

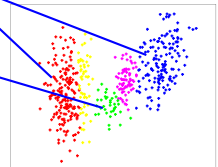
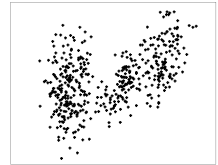
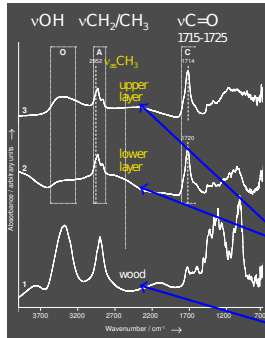
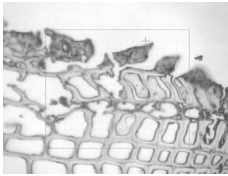


# Retour à nos violons



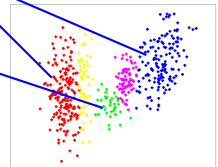
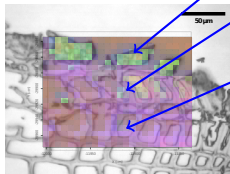
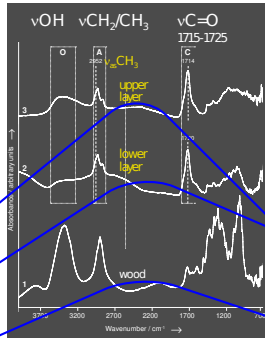
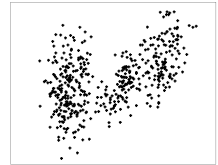
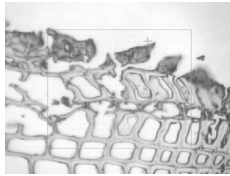
Classification

# Retour à nos violons



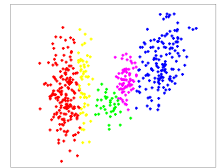
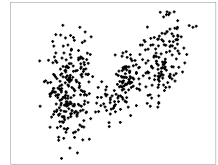
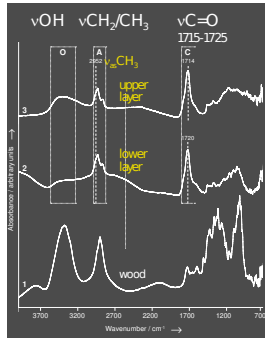
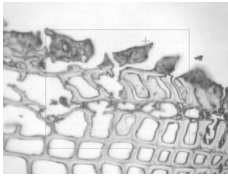
Classification

# Retour à nos violons



fo. Spatiale

# Retour à nos violons

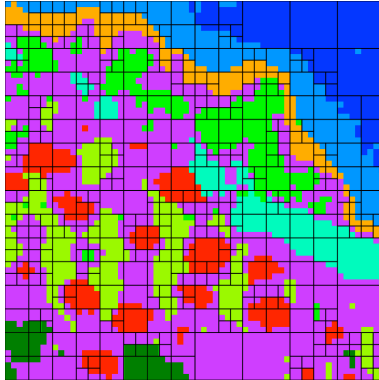


# Le secret de Stradivarius

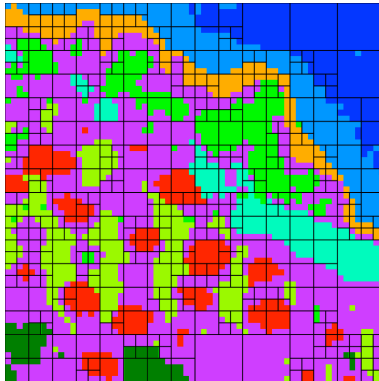




# Le secret de Stradivarius

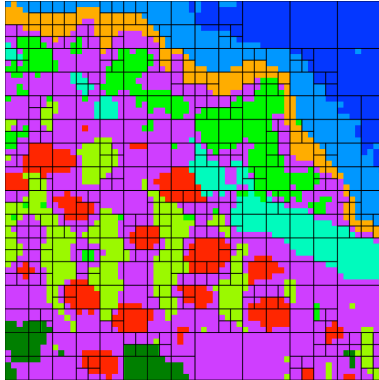


# Le secret de Stradivarius



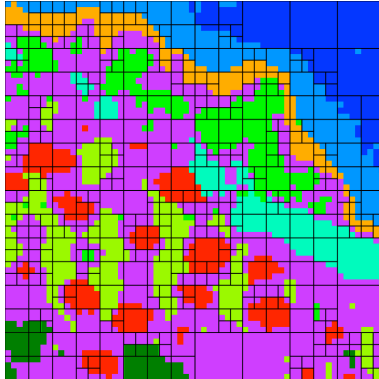
- Deux couches fines de vernis :
  - une première couche d'huile simple, similaire à celle des peintres, pénétrant légèrement le bois,
  - une seconde d'un mélange huile, résine de pin, pigments donnant cette couleur rouge caractéristique.

# Le secret de Stradivarius



- Deux couches fines de vernis :
  - une première couche d'huile simple, similaire à celle des peintres, pénétrant légèrement le bois,
  - une seconde d'un mélange huile, résine de pin, pigments donnant cette couleur rouge caractéristique.
- Technique classique pour l'époque.

# Le secret de Stradivarius



- Deux couches fines de vernis :
  - une première couche d'huile simple, similaire à celle des peintres, pénétrant légèrement le bois,
  - une seconde d'un mélange huile, résine de pin, pigments donnant cette couleur rouge caractéristique.
- Technique classique pour l'époque.
- Le secret de Stradivarius n'est pas dans le vernis !

Et après ?

# Et après ?

- Segmentation : comment prendre en compte le caractère spatial ?
  - crucial pour cette application,
  - prise en compte dans les proportions du mélange,
  - cadre théorique et algorithmique pratique.

# Et après ?

- Segmentation : comment prendre en compte le caractère spatial ?
  - crucial pour cette application,
  - prise en compte dans les proportions du mélange,
  - cadre théorique et algorithmique pratique.
- Représentation : comment trouver une représentation dans un espace de dimension raisonnable qui soit adaptée à la classification ?
  - limitation des algorithmes d'optimisation à des dimensions raisonnables,
  - point clé,
  - question ouverte !

# Et après ?

- Segmentation : comment prendre en compte le caractère spatial ?
  - crucial pour cette application,
  - prise en compte dans les proportions du mélange,
  - cadre théorique et algorithmique pratique.
- Représentation : comment trouver une représentation dans un espace de dimension raisonnable qui soit adaptée à la classification ?
  - limitation des algorithmes d'optimisation à des dimensions raisonnables,
  - point clé,
  - question ouverte !
- Classification : comment donner une réponse plus complète que des classes ?
  - classes et distance entre classes,
  - arbres "phylogéniques".



# Et après ?

- Segmentation : comment prendre en compte le caractère spatial ?
  - crucial pour cette application,
  - prise en compte dans les proportions du mélange,
  - cadre théorique et algorithmique pratique.
- Représentation : comment trouver une représentation dans un espace de dimension raisonnable qui soit adaptée à la classification ?
  - limitation des algorithmes d'optimisation à des dimensions raisonnables,
  - point clé,
  - question ouverte !
- Classification : comment donner une réponse plus complète que des classes ?
  - classes et distance entre classes,
  - arbres "phylogéniques".
- Applications à d'autres domaines.