



Détection et classification dans les signaux non stationnaires

Application à l'EMG utérin

Mohamad Khalil

Université libanaise, Faculté de génie, Branche 1
Collaboration UTT-UTC

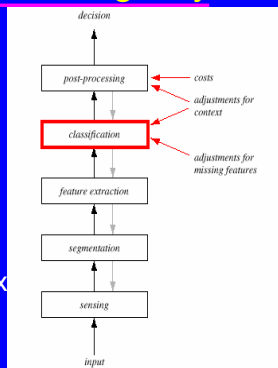
Ecole de Formation: mécatronique : Recherche et Applications
Bevrouth, du 7 au 9 février 2005

Plan

- ▾ La reconnaissance des formes
- ▾ Les paramétrages dans les signaux
- ▾ Classification supervisée
- ▾ Classification non supervisée
- ▾ Application sur les signaux EMG
 - Présentation des signaux EMG
 - Détection
 - Classification
- ▾ Autres applications

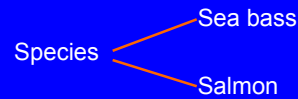
Pattern recognition: Design Cycle

- ▾ Data collection
- ▾ Feature Choice
- ▾ Model Choice
- ▾ Training
- ▾ Evaluation
- ▾ Computational Complexity



An Example

“Sorting incoming Fish on a conveyor according to species using optical sensing”



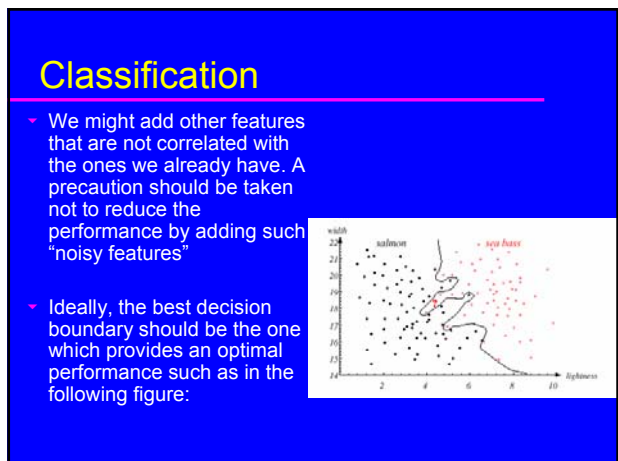
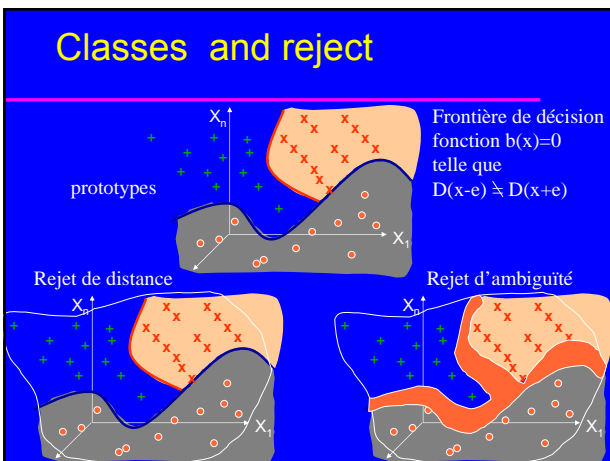
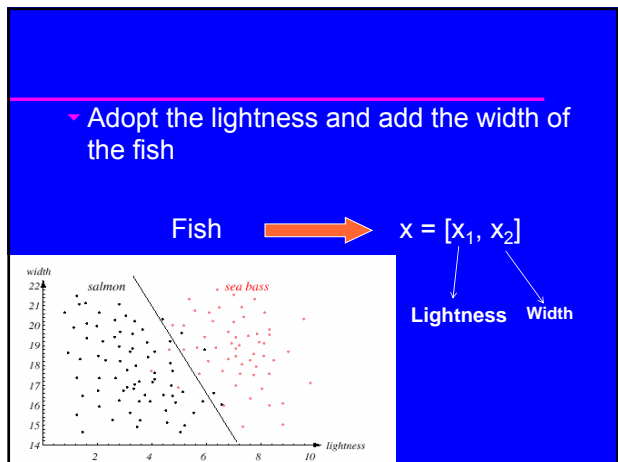
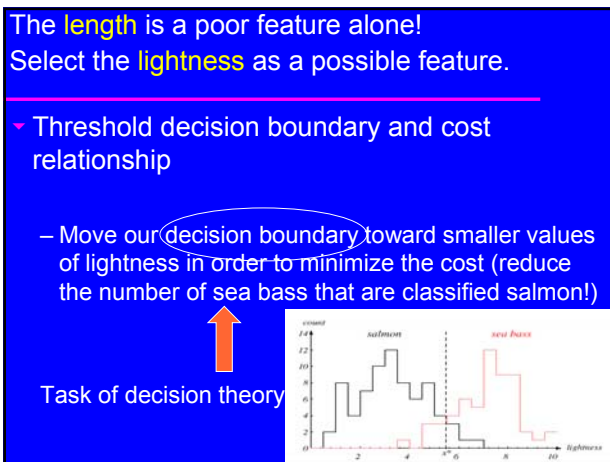
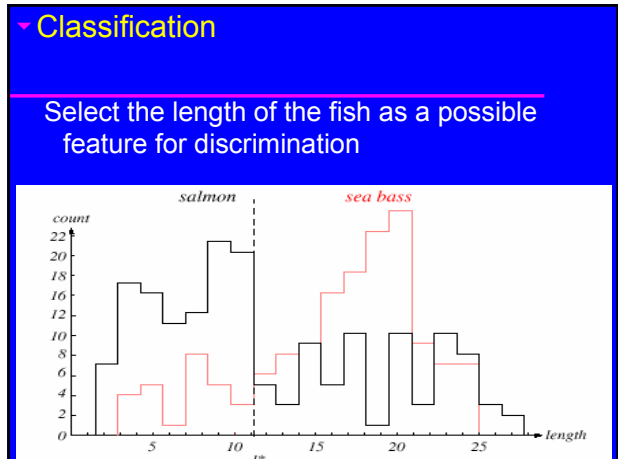
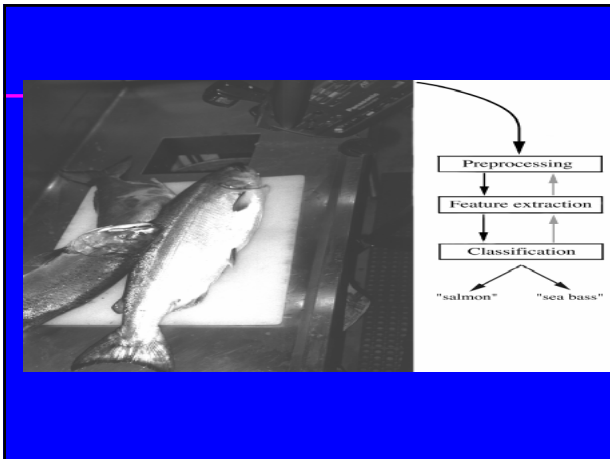
▾ Problem Analysis

- Set up a camera and take some sample images to extract features
 - » Length
 - » Lightness
 - » Width
 - » Number and shape of fins
 - » Position of the mouth, etc...

This is the set of all suggested features to explore for use in our classifier!

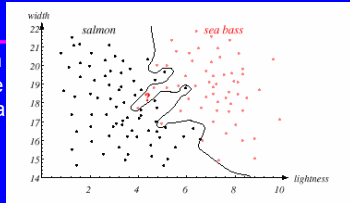
▾ Preprocessing

- Use a segmentation operation to isolate fishes from one another and from the background
- ▾ Information from a single fish is sent to a **feature extractor** whose purpose is to reduce the data by measuring certain features
- ▾ The features are passed to a **classifier**



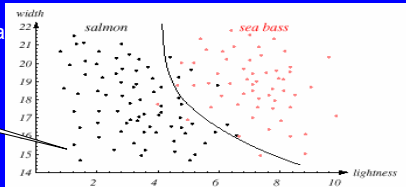
Boundaries

- However, our satisfaction is premature because the central aim of designing a classifier is to correctly classify novel input



Issue of generalization

Better?



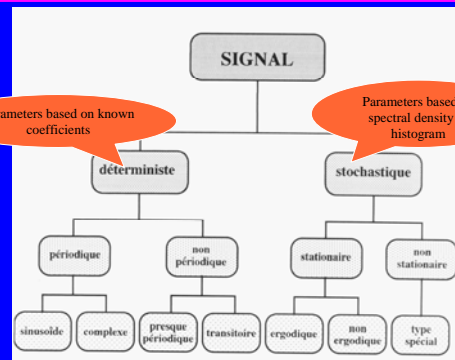
Learning and Adaptation

- Supervised learning
 - A **teacher** provides a category label or cost for each pattern in the training set
- Unsupervised learning
 - The **system** forms clusters or "natural groupings" of the input patterns

Classification and diagnostic based on signals

- Signals
 - Biomedical signals: ECG, EMG, EEG, EOG, ERG...
 - Diagnostic in machines: vibrations, sounds signals...
 - Speech signals
 - ...

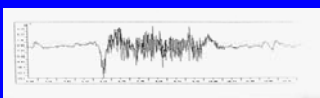
Classification of signals



Introduction



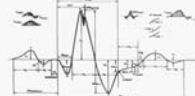
Stationnaire



Non Stationnaire

Paramétrage des signaux Quelques paramètres statistiques

• Mesures directes: min, max, durées, pentes...



• Statistiques (histogramme des amplitudes): moyenne, écart-type

• Puissance dans un intervalle de temps:

$$P(t_0) = \frac{1}{T} \int_{t_0}^{t_0+T} |s(t)|^2 dt$$

• Valeur RMS:

$$RMS(t_0) = \sqrt{P(t_0)}$$

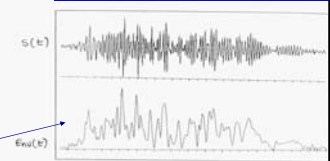
• Enveloppe:

$$Env(t) = \sqrt{|r(t)|^2 + |\hat{r}(t)|^2}$$

avec

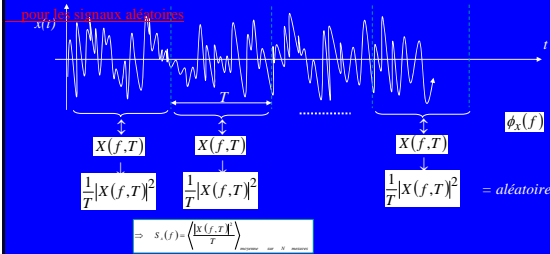
$$\hat{r}(t) = \frac{1}{j} s(t)$$

Conjugué Hilbertien



paramètres spectraux: densité spectrale

Un périodogramme correspond à une réalisation approximative de la densité de



Limitations:

- largeur et forme de la fenêtre temporelle T
- durée de la mesure \rightarrow nombre finie d'estimation de $X(f,T)$

Paramétrage des signaux: paramètres fréquentiels

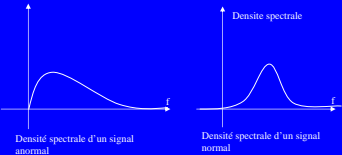
Moments spectraux

$$M_r = 2 \int_0^{\infty} f^r S_x(f) df$$

- 1- Puissance du signal : M_0
- 2- Fréquence moyenne: $MPF = M_1/M_0$
- 3- Coefficient de dissymétrie: CD

$$CD = \frac{M_3^*}{\sqrt{M_2^{3*}}}$$

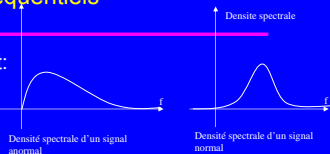
$$M_r^* = 2 \int_0^{\infty} (f - MPF)^r S_x(f) df$$



Paramétrage des signaux: paramètres fréquentiels

4- Coefficient d'aplatissement:

$$CA = \frac{M_4^*}{M_2^{2*}}$$



5- Fréquence médiane F_{med} :

Partage le spectre en deux surfaces égales

$$\int_0^{F_{med}} S_x(f) df = \int_{F_{med}}^{F_{max}} S_x(f) df$$

6- PIC de fréquence

7- Energie relative par bande de fréquence W_n

$$W_n = \frac{\int_{f_{n-1}}^{f_n} S_x(f) df}{M_0}$$

$$f_n = \frac{n}{N} f_{max} \quad \text{et} \quad 1 < n < N$$

Paramétrage des signaux: paramètres fréquentiels

8- Rapport H/L (High/Low):

$$\frac{H}{L} = \frac{\int_{f_{H1}}^{f_{H2}} S_x(f) df}{\int_{f_{L1}}^{f_{L2}} S_x(f) df}$$

$$H = [H1, H2] \quad L = [L1, L2]$$

9- Percentiles ou fractiles f_k :

$$\int_{f_{k-1}}^{f_k} S_x(f) df = k \int_0^{F_{max}} S_x(f) df$$

$$0 < k \leq 1$$

10 - Entropie spectrale H

$$H = - \int_0^{F_{max}} S_x(f) \ln[S_x(f)] df$$

Paramétrage des signaux: paramètres fréquentiels

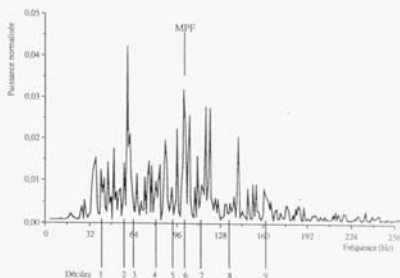


Figure 3.4: Paramétrage du spectre EMG
Les discrètes permettent de binogiser la forme liné du spectre. La MPF représente sa moyenne statistique. L'intégrer est significative de la variance spectrale et ne peut être représentée sur cette figure.

Supervised classification : Parameters and classes are known

- ▾ Probability distributions are known
 - Statistical methods: likelihood ratio
 - Bayesian networks methods
 - Neural networks bases methods
- ▾ Probability distributions are unknown,
 - Parzen methods
 - K nearest neighbors
 - Neural networks

Likelihood ratio (distributions are known):

Decide ω_1 if $P(\omega_1 | x) > P(\omega_2 | x)$; otherwise decide ω_2

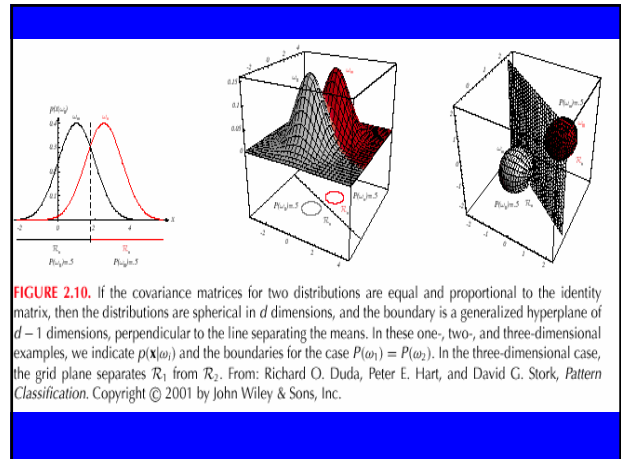
The preceding rule is equivalent to the following rule:

$$\text{if } \frac{P(x|\omega_1)}{P(x|\omega_2)} > \frac{P(\omega_2)}{P(\omega_1)}$$

Then take action α_1 (decide ω_1)

Otherwise take action α_2 (decide ω_2)

2



Unsupervised class: parameters are known, classes are unknown

- ▶ K-means method
- ▶ Fuzzy C-means algorithm
- ▶ Neural network method
 - Competitive neural network
- ▶ Principal component analysis
- ▶ Ant algorithm
- ▶ LBG algorithm: Linde, Buzo, Gray

K means algorithm (Clustering)

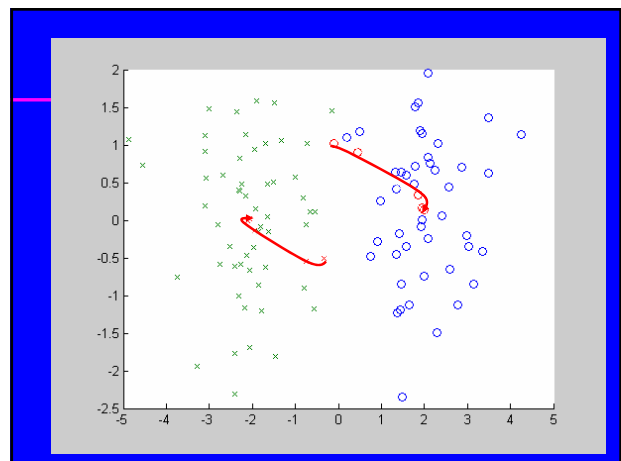
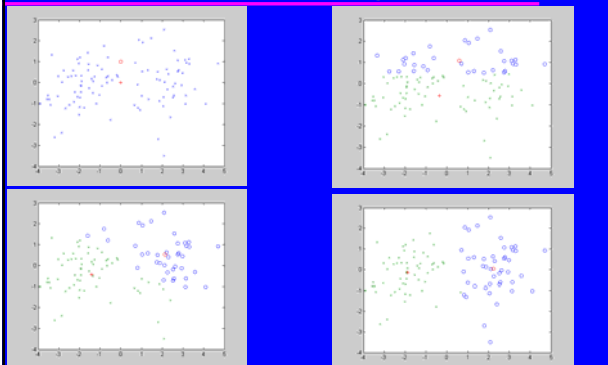
» if n is the known number of patterns and c the desired number of clusters, the k-means algorithm is:

```

Begin
  initialize  $n, c, \mu_1, \mu_2, \dots, \mu_c$ 
  (randomly selected)
  do classify  $n$  samples according to
    nearest  $\mu_i$ 
    recompute  $\mu_i$ 
  until no change in  $\mu_i$ 
  return  $\mu_1, \mu_2, \dots, \mu_c$ 
End
    
```

4

K-means method: algorithm



K- nearest neighbors (Class.)

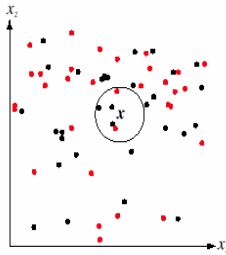
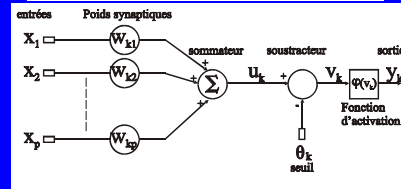
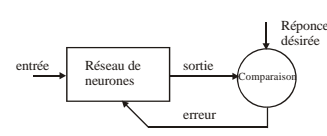
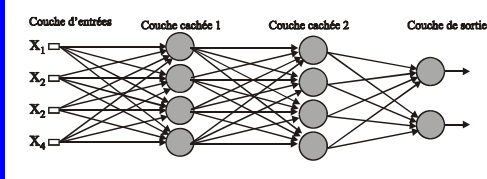


FIGURE 4.15. The k -nearest-neighbor query starts at the test point x and grows a spherical region until it encloses k training samples, and it labels the test point by a majority vote of these samples. In this $k = 5$ case, the test point x would be labeled the category of the black points. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

Neural networks

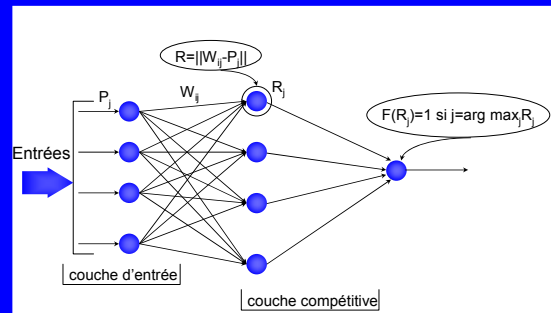


Neural networks



- Vecteur d'entrée
- Sortie désirée
- $W(t+1) = W(t) + dw$

Competitive neural network



Competitive neural network

- The adjustment of the weights will be done by:

$$\Delta_{w_{kj}} = \begin{cases} \eta(x_j - w_{kj}) & \text{if the neuron } k \text{ is winner} \\ 0 & \text{if the neuron } k \text{ is not winner} \end{cases}$$

- Initialization: Choose weight vectors $\{w_m(0); 1 \leq m \leq M\}$ randomly. Set iteration count $t = 0$.
- While Not_Converged
 - Choose the next x and compute $d(x, w_m(t)); 1 \leq m \leq M$.
 - Select $m = \text{mimmd}(x, w_m(t))$
 - $w_m(t+1) = w_m(t) + \eta(x - w_m(t))$ if m belongs to $N(m, t)$
 - $w_m(t)$ if m doesn't belong to $N(m, t)$
- End

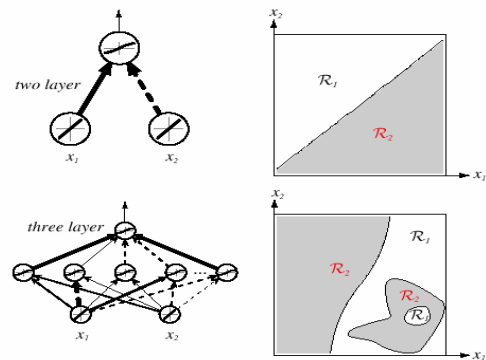


FIGURE 6.3. Whereas a two-layer network classifier can only implement a linear decision boundary, given an adequate number of hidden units, three-, four- and higher-layer networks can implement arbitrary decision boundaries. The decision regions need not be convex or simply connected. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

Application: Classification EMG

- ▶ Détection et classification dans les signaux EMG utérin
- ▶ UTC Compiègne :
Prof Catherine Marque

L'EMG utérin : introduction

- ▶ But global de l'étude
 - Détection des menaces d'accouchements prématurés

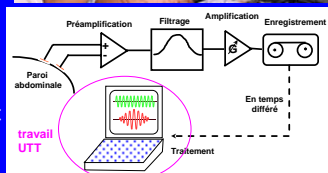
Les accouchements prématurés sont les principales causes de morbidité et de mortalité néonatale

- Développement d'un système de surveillance de grossesse
- ▶ Comment?
 - Etude de l'activité électrique utérine enregistrée à l'aide de simples électrodes de surface apposées sur l'abdomen durant la grossesse

38

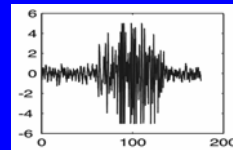
L'EMG utérin : acquisition

- ▶ Electrodes Ag-AgCl apposées sur l'axe vertical médian de la paroi abdominale en regard de l'utérus
- ▶ Diamètre des électrodes 0.5 cm
- ▶ Signaux filtrés et amplifiés entre 0.05 et 10 Hz

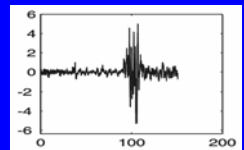


L'EMG utérin : contenu

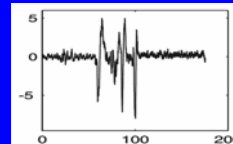
Contraction utérine



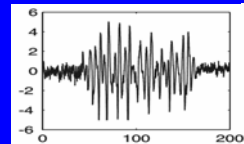
Onde d'Alvarez



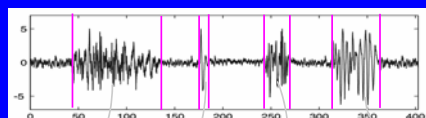
Mouvement actif foetal



Onde LDBF



L'EMG utérin : objectifs du travail



Contraction MAF Onde d'Alvarez Onde LDBF

- Détection des événements
- Identification des événements

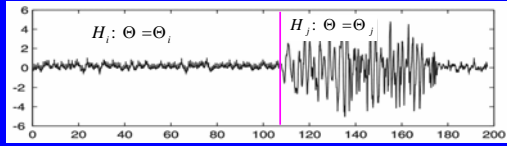
41

L'EMG utérin : le problème

- ▶ Caractéristiques des signaux
 - Le signal change d'une femme à l'autre
 - Le signal change en fonction du terme de grossesse
- ▶ Conséquence
 - Impossible de trouver une base de données unique qui représente l'EMG utérin de toutes les femmes et à chaque terme de grossesse
- ▶ Contraintes induites
 - Méconnaissance de l'espace des hypothèses
 - Méconnaissance du nombre des hypothèses
 - Non stationnarité des hypothèses

42

L'EMG utérin : formulation générale du problème



Paramètres d'hypothèses :

$$\begin{aligned} H_i: & \Theta = \Theta_i \text{ estimé} \\ H_j: & \Theta = \Theta_j \text{ inconnu} \end{aligned}$$

- Maximum de vraisemblance (GLR)
- BASSEVILLE et NIKIFOROV, LAI,...

43

La détection: Test d'hypothèse

Pour un problème de détection, il s'agit de choisir l'une des 2 hypothèses suivantes :

$$\begin{aligned} H_0: x(t) = b(t) & \text{ bruit de fond ou absence d'un signal} \\ H_1: x(t) = s(t) + b(t) & \text{ présence du signal utile} \end{aligned}$$

On définit les 4 situations suivantes :

- Non-détection correcte : choisir H0 alors que c'est H0 qui est vraie (D0,H0)
- Manque à la détection : choisir H0 alors que c'est H1 qui est vraie (D0,H1)
- Détection correcte : choisir H1 alors que c'est H1 qui est vraie (D1,H1)
- Fausse alarme : Choisir H1 alors que c'est H0 qui est vraie (D1,H0)

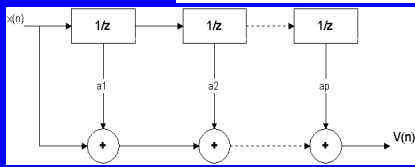
Modèle d'un signal : AR

On désire représenter un signal stochastique à l'aide d'un modèle paramétrique.

On appelle signal auto-régressif d'ordre p un signal $x(n)$ engendré par la récurrence :

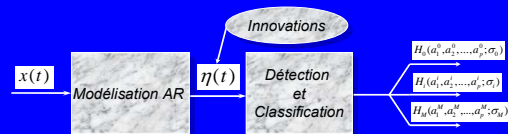


$$x(n) = -\sum_{i=1}^p a(i).x(n-i) + v(n) \quad G(z) = \frac{x(z)}{v(z)} = \frac{1}{\sum_{i=0}^p a(i).z^{-i}}; \text{ avec } a(0) = 1$$



Prétraitement

- Nécessité d'échantillons indépendants
- Modèle AR

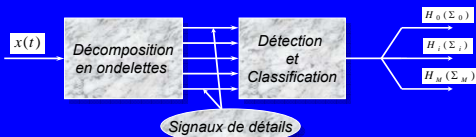


- » Les coefficients AR contiennent des informations sur le contenu fréquentiel
- » La variance porte l'information sur l'énergie

46

Prétraitement

- Ondelettes : explorer le contenu fréquentiel



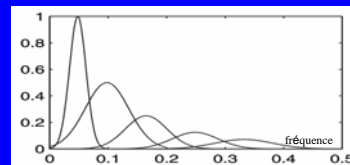
- » Chaque niveau d'échelle porte une information spécifique en fréquence

47

Prétraitement

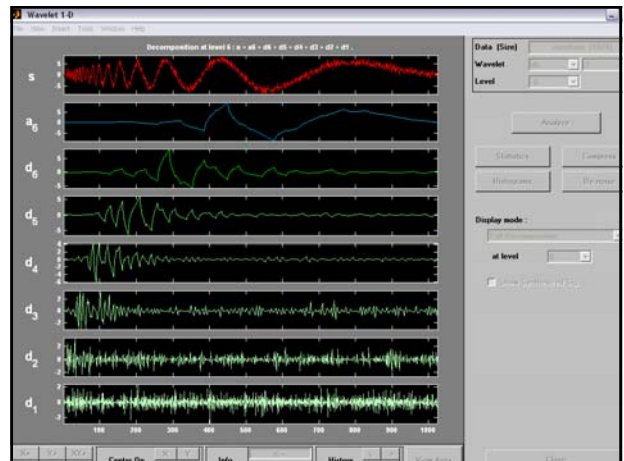
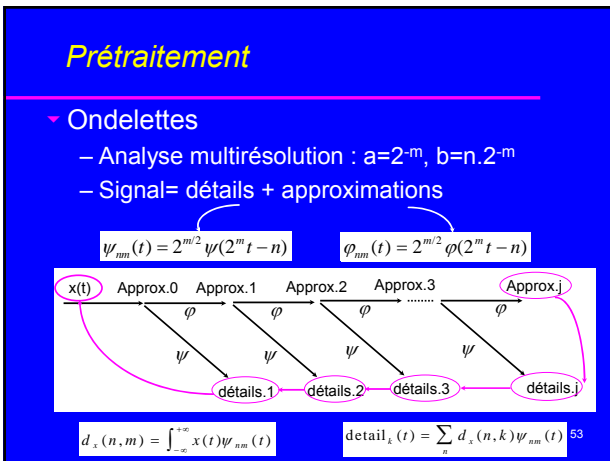
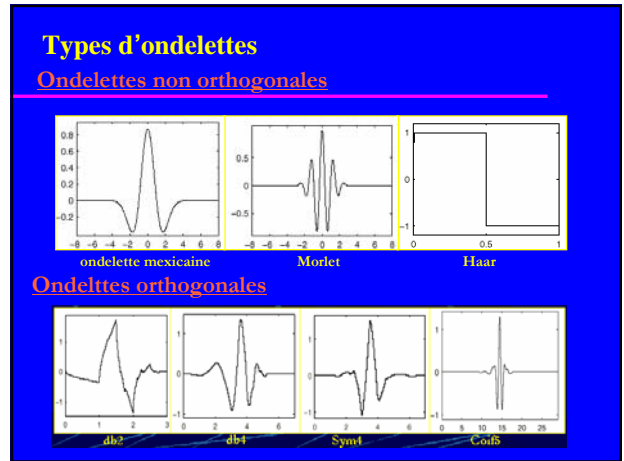
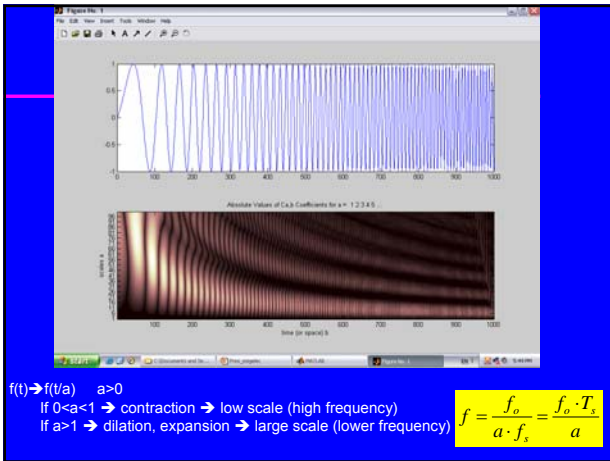
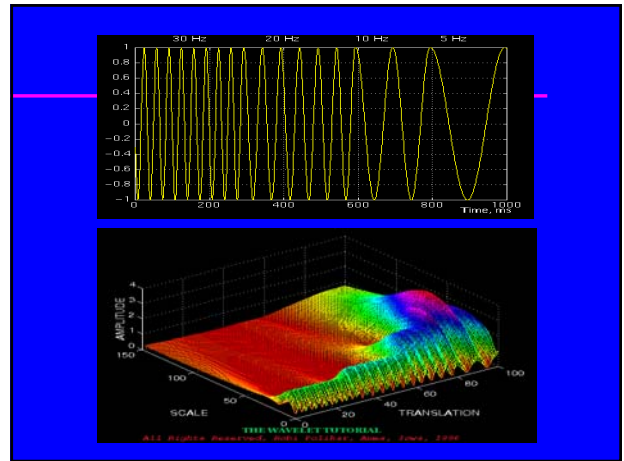
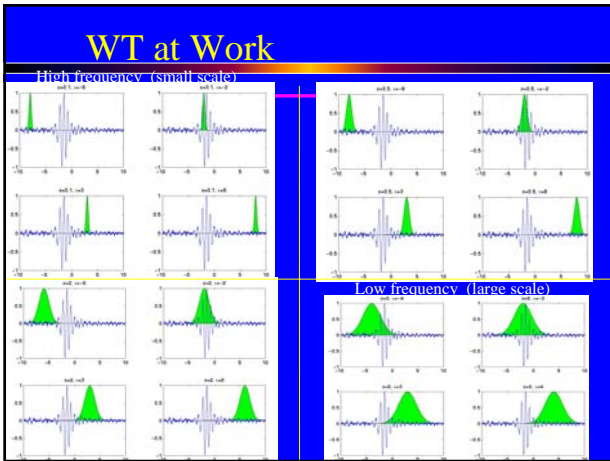
- Ondelettes : principe

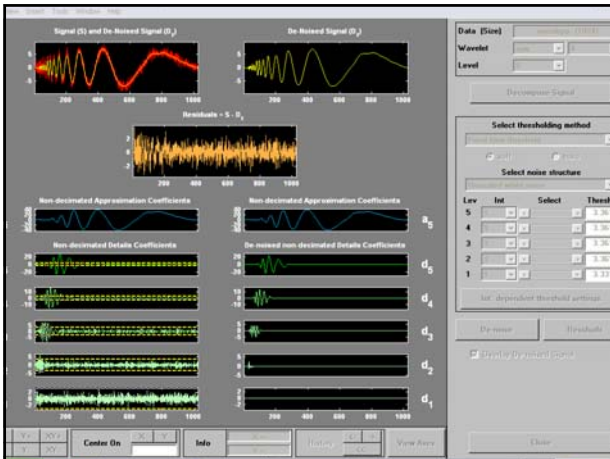
$$\begin{aligned} T_x^\psi(a, b) &= \int_{-\infty}^{+\infty} x(t) \psi_{ab}(t) dt \\ \psi_{ab}(t) &= \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \end{aligned} \quad \text{Ondelette}$$



Banc de filtres à surtension constante = $\frac{\text{center frequency}}{\text{bandwidth}}$

48





Prétraitement

- Ondelettes : exemple de décomposition

Décomposition par paquets d'ondelettes

- Décompositions successives des approximations et des détails:
 - Cet algorithme conduit à une décomposition en sous bandes de fréquences du signal.

$S = A1 + AAD3 + DAD3 + DD2$

Décomposition par paquets d'ondelettes

- Obtenir la meilleure décomposition par l'algorithme de Coifman-Wickerhauser

Algorithme de détection classique

- Somme cumulée: somme des rapports de vraisemblance

$$S_i = \sum_{j=1}^i \log \frac{f_{\theta_1}(X_j)}{f_{\theta_0}(X_j)}$$

Algorithme de détection classique

- Somme cumulée: somme des rapports de vraisemblance

$$S_i = \sum_{j=1}^i \log \frac{f_{\theta_1}(x_j)}{f_{\theta_0}(x_j)}$$

Instants d'arrêts

$$t_1 = \inf \left\{ n \geq 1 : S_n - \min_{0 \leq k \leq n} S_k \geq h \right\}$$

Algorithme de détection classique

- Somme cumulée: somme des rapports de vraisemblance

$$S_t = \sum_{i=1}^t \log \frac{f_{\theta_1}(x_i)}{f_{\theta_0}(x_i)}$$

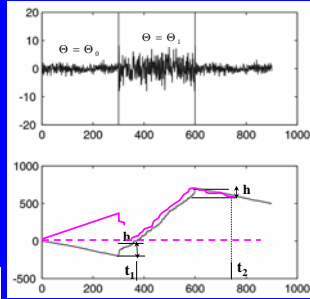
Instants d'arrêts

$$t_1 = \inf \left\{ n \geq 1 : S_n - \min_{0 \leq r \leq n} S_r \geq h \right\}$$

$$t_2 = \inf \left\{ n \geq t_1 : \max_{t_1 \leq r \leq n} S_r - S_n \geq h \right\}$$

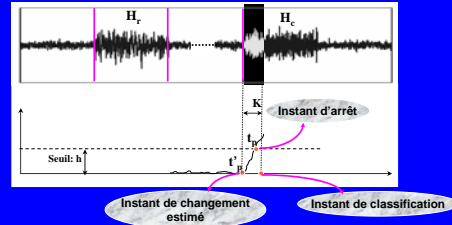
Equation généralisée:

$$t_p = \inf \left\{ n \geq t_{p-1} : \max_{t_{p-1} \leq r \leq n} |S_r - S_n| \geq h \right\}$$



Algorithme de détection classique

- Cas général: classification



Classification: hypothèse la plus proche
Minimum des sommes cumulées

$$H_{pp} : pp = \text{Arg} \left[\min_{r=1..M} S'_r \right] \quad S'_r = \sum_{i=t_{r-1}}^t \log \frac{f_{\theta_r}(x_i)}{f_{\theta_0}(x_i)}, \quad r=1..M$$

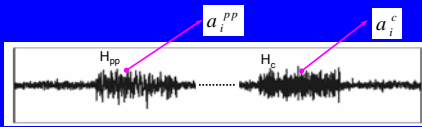
62

Algorithme de détection classique

- Modélisation AR

» Paramètre: $\Theta_i = (a_i^t, a_i^c, \sigma_i)$

» Rapport de vraisemblance: $S'_r = \sum_{i=N_{r-1}}^t \log \frac{f_{\theta_r}(x_i / \mathbf{X}_{i-1}^r)}{f_{\theta_0}(x_i / \mathbf{X}_{i-1}^0)}, \quad r=1..M$



Test de rejet: Distribution de l'erreur de prédiction

$$f^{pp}(\text{erreur}) = \left[\frac{1}{\sqrt{2\pi\sigma_{pp}}} \exp\left(-\frac{\text{erreur}^2}{2\sigma_{pp}^2}\right) \right]^k \quad \text{erreur} = \frac{1}{k} \sum_{i=1}^k (\eta_i^{pp})^2$$

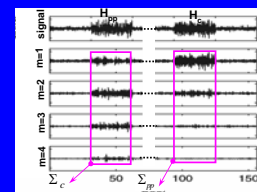
63

Algorithme de détection classique

- Test de rejet: Ondelettes

» Paramètre: $\Theta_i = \Sigma_i$

» Rapport de vraisemblance: $S'_r = \sum_{i=N_{r-1}}^t \log \frac{f_{\theta_r}(X_i)}{f_{\theta_0}(X_i)}, \quad r=1..M$



Test de rejet :

$$H_0 : \Sigma_c = \Sigma_{pp} \\ H_1 : \Sigma_c \neq \Sigma_{pp}$$

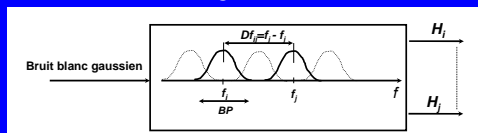
Test de Box

$$\left(1 - \frac{2d^2 + 3d - 1}{6d + 6} \left(\sum_{i=1}^k \frac{1}{n_i} - \frac{1}{n-2} \right) \right) \left((n-2) \text{Log} \left| \frac{n-2}{n-2} W \right| - \sum_{i \in \{c,s\}} (n_i - 1) \text{Log} |\Sigma_i| \right) \rightarrow \chi_{d(d+1)}^2$$

64

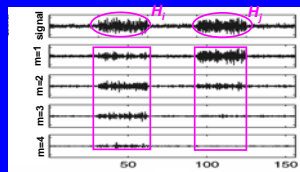
Algorithme de détection classique

- Performances : signaux simulés



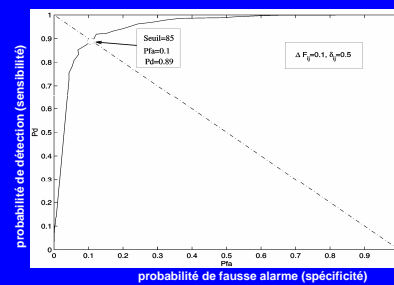
Distance entre deux hypothèses:

$$\delta_{ij} = \frac{f_i - f_j}{BP}$$



Algorithme de détection classique

- Performances de la détection : courbe COR



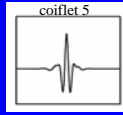
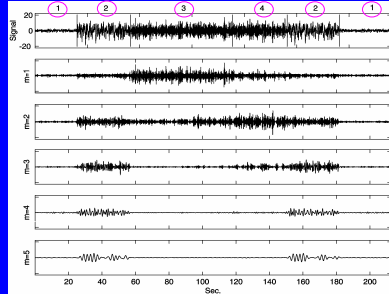
500 segments utilisés

66

Algorithme de détection classique

Performances sur la détection classification

Signaux simulés



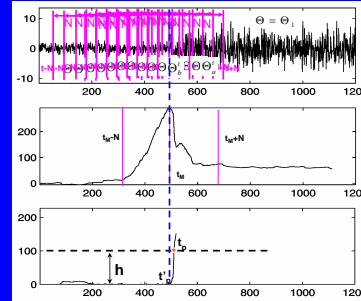
seuil de la détection: 85

Pfa de la classification: 0.01

67

Approche dynamique pour la détection

Principe



Hypothèses instantanées

$$H_0^i: X_i; i = \{t-N, \dots, t-1\} \rightarrow f_{\Theta_0}$$

$$H_1^i: X_i; i = \{t, \dots, t+N-1\} \rightarrow f_{\Theta_1}$$

Somme cumulée dynamique

$$DCS(H_1^i, H_0^i) = \sum_{t=N}^t \log \frac{f_{\Theta_1}(X_i)}{f_{\Theta_0}(X_i)}$$

Fonction de détection

$$g(t) = \max_{t_p, t_n} [DCS(H_1^i, H_0^i)] - DCS(H_1^i, H_0^i)$$

Instant d'arrêt

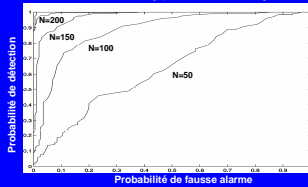
$$t_p = \inf \{n > t_{p-1}; g(n) > h\}$$

68

Approche dynamique pour la détection

Effet de la taille N (fenêtres égales)

- » durée des événements courts à détecter
- » le retard global à la détection
- » le nombre de points utilisés pour estimer les paramètres des hypothèses dynamiques



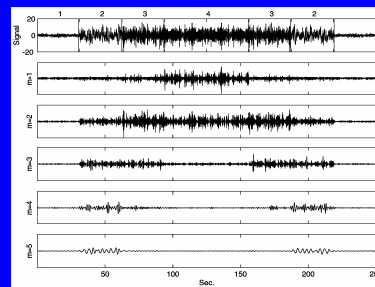
Choix : N=200

69

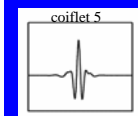
Performance de la détection en fonction de N

Approche dynamique pour la détection

Résultats sur la détection classification



Ondelette:



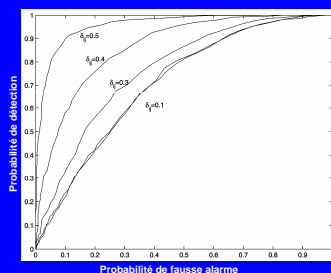
seuil de la détection : 35

Pfa de la classification : 0.01

70

Approche dynamique pour la détection

Performances de la détection : courbes COR



N=200, 1000 segments utilisés

71

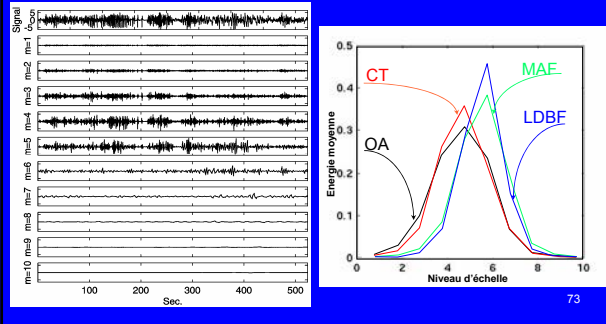
Adaptation et application aux signaux réels

- » Choix de l'ondelette
- » Choix des échelles
- » Choix du seuil de détection
- » Choix du seuil de classification

72

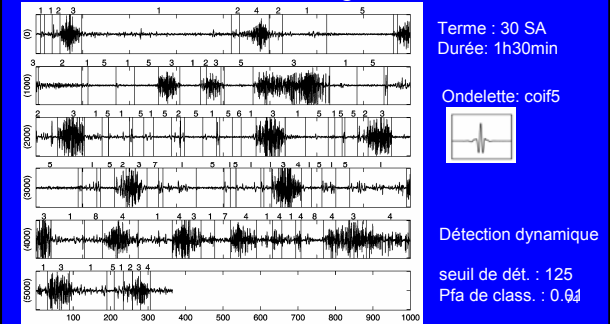
Adaptation et application aux signaux réels

Choix de l'ondelette et des échelles



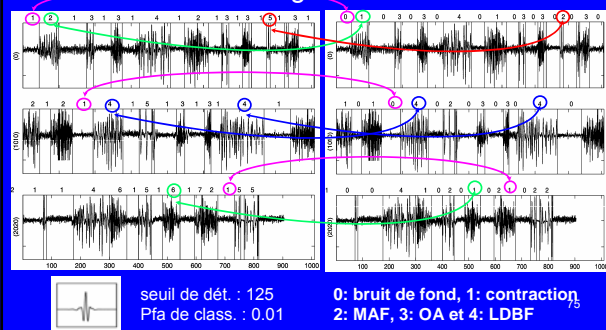
Adaptation et application aux signaux réels

Résultats sur des enregistrements EMG

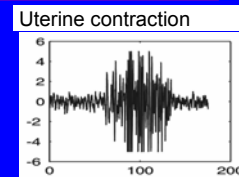


Identification des groupes par réseau de neurones

Résultats sur le signal réel: identification

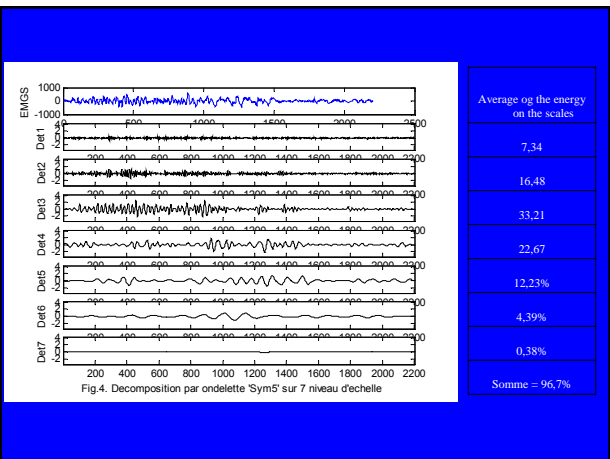
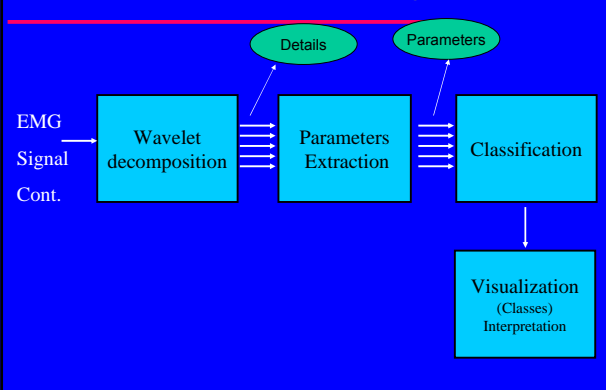


Uterine contractions



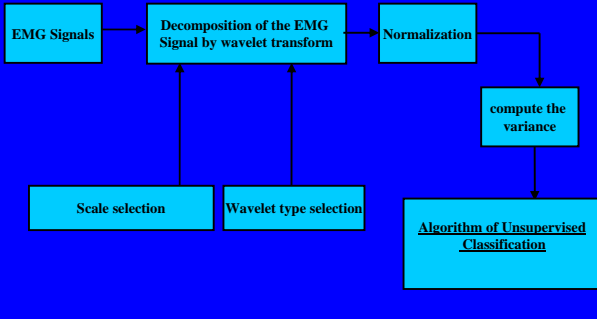
- Characteristics
 - Length
 - Inter and intra frequency
 - Frequency content: Shifting in frequency according to the duration of pregnancy
- Method: Wavelet transform

General Block Diagram



Parameters Extraction

- 5 parameters containing variances



Classification algorithm

1. Initialize the center of the first class with the first vector.
2. Take a new vector, find the nearest class center.
3. Compare each component of the new vector to the classes using the Fisher test
4. If this vector is identical to the class centre, we adapt each class
5. otherwise we create a new class
6. Goto 2

$$F_{n_1-1, n_2-2} = \frac{n_1 S_1^2}{n_1 - 1} \frac{n_2 - 1}{n_2 S_2^2}$$

Principle Component Analysis

- After classification, a principal component projection is used to evidence the result of our classification.
- This method reduces the dimension of the representation space to keep only the most important information, by computing the estimators of the principle axis of the ellipsoid, characteristic of a multi-dimensional Gauss's distribution

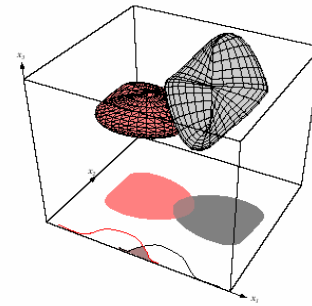
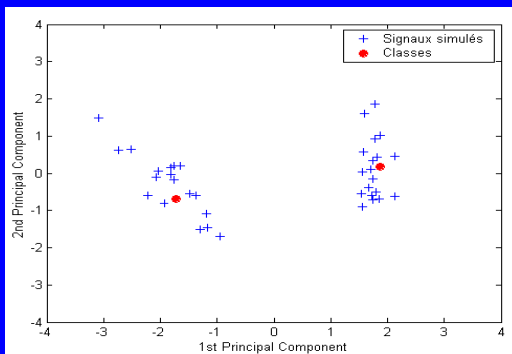
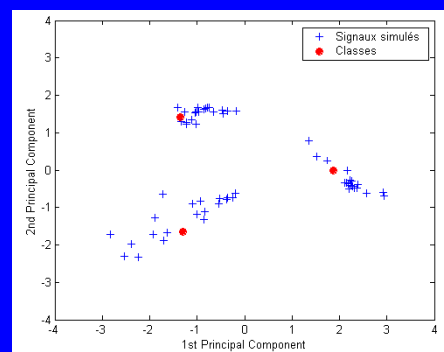


FIGURE 3.3. Two three-dimensional distributions have nonoverlapping densities, and thus in three dimensions the Bayes error vanishes. When projected to a subspace—here, the two-dimensional $x_1 - x_2$ subspace or a one-dimensional x_1 subspace—there can be greater overlap of the projected distributions, and hence greater Bayes error. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

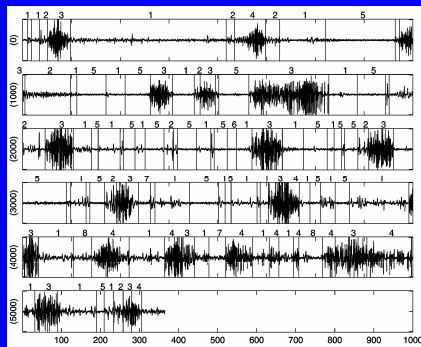
Classification on Simulated Signals 2 classes



Result on simulated signals : 3 classes

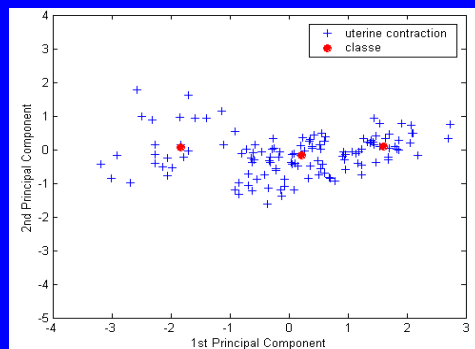


Real Uterine EMG signals

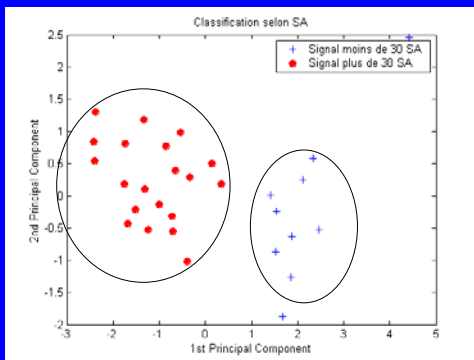


Pregnancy: 30 Weeks
Duration: 1h30min

Real Signals classification: 3 classes



Interpretation according to weeks of gestation



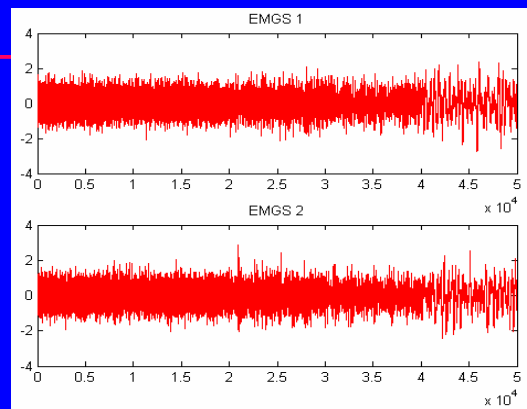
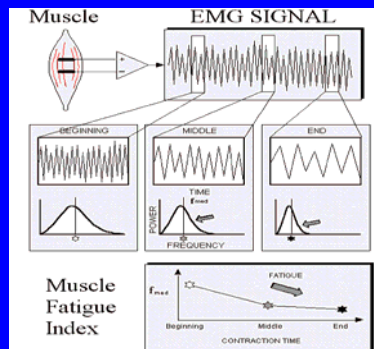
Autres applications

- Etude des signaux musculaires pour la détection du siège de la voiture le plus confortable
- UTT (Troyes) Prof. Jacques Duchene
- Collaborations: Renault

Les signaux physiologiques:EMG

• Traitement:

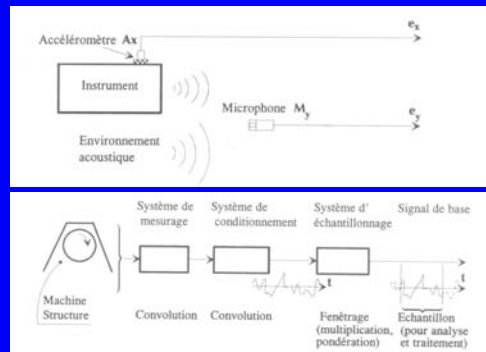
Analyse de la fatigue :
Le signal change sa fréquence en fonction de la fatigue



Autres applications

- Diagnostic des pannes sur un système mécanique par reconnaissance de formes
- UL, faculté de génie Branche 1
Dr Khaled KHALIL
- ESIGELEC de Rouen et UL
- Financement CNRS

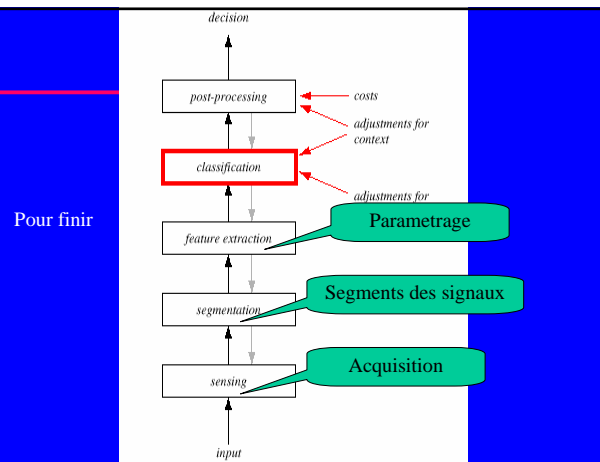
Chaîne d'acquisition



Paramétrage

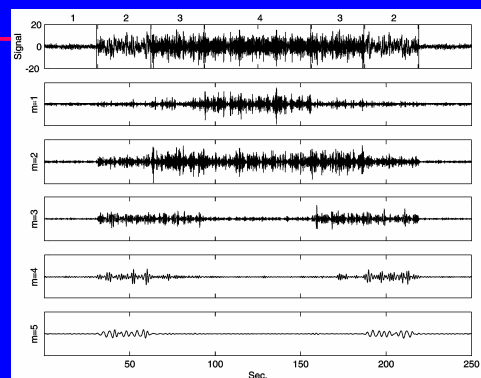
- Kurtosis: Moment statique d'ordre 4 centré et normalisé. Aplatissement de la distribution.
- Kurtosis par bande de fréquence.
- Décalage en fréquence avec l'usure des engrenages,
- Nouveaux composants « fantômes » apparaissent dans les spectres

Paramètres fréquentiels: Energie après décomposition en ondelettes.



Perspectives

- travail de thèse en cours :
 - Algorithme de détection, classification séquentiel basé sur la théorie des ondelettes et les algorithmes séquentiels de détection (DCS+Ondelettes)
 - Appliqué sur n'importe quel type de signaux présentant un changement fréquentiel et/ou amplitude.



Perspectives

- Thèse 2:
 - Classification des contractions utérines en fonction du temps d'accouchement
 - Paramétrage: temporels et fréquentiels
 - Classification non supervisée
 - Réseaux de neurones + ondelettes
 - ...