



HAL
open science

Détection et classification des signaux non stationnaires par utilisation des ondelettes. Application aux signaux électromyographiques utérins

Marwa Chendeb

► **To cite this version:**

Marwa Chendeb. Détection et classification des signaux non stationnaires par utilisation des ondelettes. Application aux signaux électromyographiques utérins. Traitement du signal et de l'image [eess.SP]. Université de Technologie de Troyes, 2006. Français. NNT : . tel-00452410

HAL Id: tel-00452410

<https://theses.hal.science/tel-00452410>

Submitted on 2 Feb 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

pour l'obtention du grade de

**DOCTEUR de l'UNIVERSITE
DE TECHNOLOGIE DE TROYES
Spécialité : OPTIMISATION ET SURETE DES SYSTEMES**

présentée et soutenue par

Marwa CHENDEB

le 14 mars 2006

**Détection et classification des signaux non stationnaires par
utilisation des ondelettes. Application aux signaux
électromyographiques utérins**

JURY

Mr C. RICHARD	PROFESSEUR DES UNIVERSITES	Président
Mr G. CARRAULT	PROFESSEUR DES UNIVERSITES	Rapporteur
Mr J. DUCHENE	PROFESSEUR DES UNIVERSITES	Directeur de thèse
Mr M. KHALIL	ENSEIGNANT CHERCHEUR - HDR	Directeur de thèse
Mme C. MARQUE	PROFESSEUR DES UNIVERSITES	Rapporteur

A mon Père et ma Mère

Aucune dédicace ne pourra traduire mon profond amour et gratitude pour l'affection que vous m'avez offerte. Les sacrifices illimités, encouragements et prières ont été pour moi source de zèle et gage de réussite.

Que cette thèse soit l'expression de ma profonde tendresse et reconnaissance. J'espère être digne et à la hauteur de votre confiance et de vos espérances. Je prie Allah pour vous jouir de bonne santé toute la vie et que vos vœux soient exaucés.

A mes chers frères et sœurs

Que cette thèse soit un témoignage de la profondeur des sentiments d'amour, de considération et de respect que j'éprouve à vos égards. Je prie Allah tout puissant de vous protéger et de vous accorder réussite et bonheur.

A Salah la lumière de ma vie

Que ce modeste travail puisse traduire la profondeur de l'amour et de la reconnaissance infinie que je porte pour toi.

A ma belle famille

Permettez-moi de vous exprimer ma gratitude, la vraie reconnaissance et la plus profonde considération.

A mes meilleurs amis Rana, Nadjat, Valérie, Yasmine, Rima, Hassan, Georges, Paul, Bassem...

A tous mes amis

Que vous trouvez dans ce travail l'expression de ma sincère reconnaissance pour votre disponibilité et votre gentillesse.

A mes enseignants à la faculté de génie de l'Université Libanaise

Je vous saurais gré du savoir que vous m'avez transmis des connaissances que vous m'avez prodiguées. Que ce travail soit le témoignage de ma gratitude et de ma grande estime.

Remerciements

A mon directeur de thèse Mr Jacques DUCHÈNE

Je suis profondément touchée par le grand honneur que vous me faites en acceptant de diriger cette thèse. Permettez-moi de vous manifester ma grande admiration pour votre compétence, vos qualités humaines : modestie, bienveillance et moralité qui resteront pour moi un exemple à suivre dans ma vie professionnelle. Vous m'avez toujours accueilli avec amabilité et simplicité habituelles malgré vos nombreuses occupations. Vous avez veillé sur la réalisation de cette thèse et vous m'avez mené à bon escient. Sans votre aide précieuse, vos suggestions, vos conseils judicieux et vos encouragements incessants, ce travail n'aurait pu être accompli. Puisse ce travail vous satisfaire et témoigner de ma grande reconnaissance et ma profonde estime.

A mon co-directeur de thèse Mr Mohamad KHALIL

Vous m'avez encouragé à faire cette thèse et vous m'avez aidé à surmonter beaucoup de difficultés rencontrées grâce à votre enthousiasme et votre dynamisme. Veuillez trouver dans ce travail l'expression de mes sincères reconnaissances et respects.

A Mr Cédric RICHARD, le président de Jury,

Vous me faites un grand honneur en acceptant de siéger au jury de cette thèse.

A Mme Catherine MARQUE

Je suis très honorée de vous compter parmi les rapporteurs et je vous remercie vivement d'avoir accepté de juger ce travail. Permettez-moi de vous manifester ma grande admiration pour votre courtoisie et vos conseils. Veuillez trouver dans ce travail, l'assurance de ma gratitude et de mon respect.

A Mr Guy CARRAULT

Ravie et fière de l'honneur que vous m'avez fait en acceptant aimablement d'être un rapporteur de ce travail. Puisse ce travail vous satisfaire et témoigner de ma haute considération et ma profonde estime.

A Mr David HEWSON, Mme Babeth DUCHÈNE, Mme Leïla BOULAHIA

Je vous remercie pour vos aides. Veuillez trouver ici l'expression de mes grands estimes et remerciements.

A tous les membres du laboratoire et les personnels de l'Université de Technologie de Troyes

A Mr Jérémy TERRIEN

A Mr Haïssam ZIADÉ mon enseignant à la faculté de génie de l'université libanaise

Vos qualités humaines et votre modestie ont toujours suscité mon admiration. Que ce travail soit le témoignage de ma gratitude et de ma grande estime.

Table des matières

Introduction	9
Chapitre 1 - Position du problème.....	15
1.1 Introduction	15
1.2 Objectifs détaillés de l'étude	15
1.3 Description des signaux	18
1.3.1 Signal EMG.....	18
1.3.2 Les différents événements contenus dans l'EMG	19
1.4 Modélisation de l'EMG.....	24
1.5 Acquisition et Prétraitement de l'EMG utérin	26
1.6 Approches antérieures de traitement de l'EMG utérin.....	27
1.7 Différentes étapes du traitement des signaux EMG	28
1.8 Synthèse	30
Chapitre 2 - Représentations à base d'ondelettes et de paquets d'ondelettes	33
2.1 Introduction	33
2.2 Pourquoi les décompositions en ondelettes et en paquets d'ondelettes ?	34
2.3 Les bases de la décomposition discrète en ondelettes.....	36
2.3.1 Analyse multirésolution	37
2.3.2 Filtres miroirs conjugués.....	38
2.3.3 Propriétés et caractéristiques des ondelettes orthogonales.....	40
2.3.4 Décomposition en ondelettes orthogonales : Algorithme pyramidal de Mallat.....	41
2.4 Les paquets d'ondelettes	46
2.4.1 Décomposition en paquets d'ondelettes.....	46
2.4.2 Bancs de filtres de paquets d'ondelettes	49
2.4.3 Calcul des coefficients des paquets d'ondelettes	50
2.4.4 Arbre admissible	52
2.5 Sélection d'une meilleure base des paquets d'ondelettes.....	53
2.5.1 Les algorithmes fondamentaux de recherche de la meilleure base	54
2.5.2 Les critères de sélection	55
2.5.3 Algorithme de Coifman Wicherhauser	57

2.5.4	Sélection d'une meilleure base pour la compression [Coif92]	57
2.5.5	Sélection d'une meilleure base pour le débruitage.....	58
2.5.6	Sélection d'une meilleure base pour la classification supervisée	58
2.5.7	Sélection de la meilleure base pour la détection des ruptures fréquentielles	59
2.6	Conclusion.....	61

Chapitre 3 - Sélection d'une meilleure base pour la détection des transitoires..... 63

3.1	Introduction	63
3.2	Critère de sélection.....	64
3.3	Densité gaussienne généralisée pour les coefficients des paquets d'ondelettes.....	68
3.3.1	Estimation des paramètres d'échelle et de forme α et β	68
3.3.2	Distribution des coefficients des paquets d'ondelettes dans le cas de l'EMG utérin	70
3.4	Estimation de la distance de Kullback Leibler.....	75
3.5	Estimation de la distribution de \hat{K}	77
3.5.1	Distribution obtenue dans le cas où il n'y a pas d'événements dans les enregistrements ($\alpha_i = \alpha_j$)	78
3.5.2	Cas où des événements sont présents dans les enregistrements ($\alpha_i \neq \alpha_j$)	80
3.6	Sélection des paquets d'ondelettes	81
3.6.1	Caractérisation des paquets d'ondelettes pour la détection.....	81
3.6.2	Effet de la longueur de la fenêtre N	82
3.6.3	Construction de l'algorithme de sélection de la meilleure base.....	83
3.6.4	Paramètres attachés à l'algorithme de sélection de la meilleure base.....	86
3.7	Application de la méthode de sélection de la meilleure base.....	87
3.7.1	Résultats de la méthode appliquée sur les signaux synthétiques	87
3.7.2	Résultats de la méthode appliquée sur les signaux réels.....	89
3.8	Discussion et Conclusion	91

Chapitre 4 - Détection et segmentation de l'EMG utérin..... 95

4.1	Introduction	95
4.2	Utilisation des coefficients d'ondelettes pour la détection.....	97
4.3	Théorie du décalage des coefficients d'ondelettes	98
4.3.1	Support temporel des coefficients d'ondelettes	98
4.3.2	Cas de la décomposition discrète en ondelettes	100
4.3.3	Cas de la décomposition en paquets d'ondelettes	103
4.3.4	Synthèse	106
4.4	Algorithmes de détection de changements et application aux coefficients d'ondelettes	106

4.4.1 Méthodes de détection de ruptures dans le cas d'un changement à paramètres connus.....	107
4.4.2 Méthodes de détection des ruptures dans le cas de paramètres inconnus	110
4.4.3 Fusion des instants de changement	113
4.5 Validation et performance des méthodes	115
4.5.1 Performance de la correction de décalage.....	115
4.5.2 Performance des algorithmes de détection sur les meilleurs paquets choisis après correction et fusion des instants de changement	122
4.6 Performance liée à l'utilisation directe des coefficients d'ondelettes	127
4.7 Détection en temps réel	128
4.8 Conclusion.....	132
Chapitre 5 - Classification et identification des événements de l'EMG utérin.....	135
5.1 Introduction	135
5.2 Reconnaissance des Formes et paquets d'ondelettes	137
5.3 Sélection de la meilleure base pour la discrimination entre classes.....	138
5.3.1 Critère de discrimination entre classes pour la sélection de la meilleure base	138
5.3.2 Sélection de la meilleure base pour la classification.....	142
5.4 Méthodes de classification	143
5.4.1 Règle du plus proche voisin	143
5.4.2 Distance de Mahalanobis et classification	144
5.4.3 Réseaux de Neurones : Perceptron multicouches	145
5.4.4 Machine à Vecteurs Supports.....	146
5.5 Application aux événements de l'EMG utérin.....	151
5.5.1 Sélection des paquets discriminants.....	152
5.5.2 Vérification de la performance du choix des paquets : utilisation de la distance de Mahalanobis	153
5.5.3 Vérification de la performance du choix des paquets : utilisation de la méthode des K plus Proches Voisins	156
5.5.4 Conclusion.....	156
5.6 Classification à partir des paquets sélectionnés	157
5.6.1 Distance de Mahalanobis et des K Plus Proches Voisins.....	158
5.6.2 Réseaux de neurones	158
5.6.3. Machines à Vecteurs Supports	159
5.7 Conclusion.....	160
Conclusion générale	163
Annexe A	167
Annexe B	171

Annexe C	175
Annexe D	179
Annexe E	183
Annexe F	185
Bibliographie	191

Abréviations

AP	Accouchement Prématuro
Coifx	Ondelette Coiflet no. x
CUSUM	Cumulative sum
dbx	Ondelette daubechies no. x
DCS	Dynamic Cumulative Sum
DSP	Densité Spectrale de Puissance
DWT	Discrete Wavelet Transform
EHG	Electrohystérogramme (signal EMG utérin)
EMG	Electromyogramme
FWH	Fast Wave High
FWL	Fast Wave Low
GGD	Generalized Gaussian Density
K-L	Kullback Leibler
KPPV	K Plus Proches Voisins
LDBF	Longue Durée Basse Fréquence
MAF	Mouvement Actif Foetal
MAP	Menace d'Accouchement Prématuro
NF	Noeud Fils
NP	Nœud Père
OMS	Organisation Mondiale de la Santé
SA	Semaine d'Aménorrhée
SVM	Support Vector Machine
Symx	Ondelette Symlet no. x
WPT	Wavelet Packet Transform

Notations

α	Paramètre d'échelle de la densité gaussienne généralisée
$\hat{\alpha}$	Paramètre d'échelle estimée de la densité gaussienne généralisée
β	Paramètre de forme de la densité gaussienne généralisée
λ	Paramètre de la loi exponentielle
$\psi_{j,n}$	Ondelette analysante (transformation discrète)
$\phi_{j,n}$	Fonction d'échelle (transformation discrète)
$\Gamma(z)$	Fonction gamma
$\Pi(z)$	Fonction digamma
$a_j(k)$	$k^{\text{ième}}$ coefficient d'approximation à l'échelle j
$C_{j,n}(k)$	$k^{\text{ième}}$ coefficient du paquet d'ondelettes à l'échelle j et au paquet numéro n
D_{\max}	Statistique du test de Kolmogorov Smirnov
$d_j(k)$	$k^{\text{ième}}$ coefficient de détail à l'échelle j
f_{θ_i}	Densité de probabilité de paramètre θ_i
g	Filtre passe haut (transformation en ondelettes)
h	Filtre passe bas (transformation en ondelettes)
j	Échelle courante de la décomposition en paquets d'ondelettes
J	Niveau maximal de décomposition
K	Distance de Kullback - Leibler
\hat{K}	Distance estimée de Kullback Leibler
L	Longueur des réponses impulsionnelles des filtres correspondant à une ondelette.
M	Nombre de classes
N	Longueur de la fenêtre
t_b''	Position du barycentre énergétique
$t_j(k)$	Instant corrigé d'un coefficient de détail ou d'approximation détecté à l'échelle j

$t_{j,n}(k)$	Instant corrigé d'un coefficient d'un paquet n détecté à l'échelle j
V_j	Espace d'approximation à l'échelle j
W_j	Espace de détail à l'échelle j
$W_{j,n}$	Espace du paquet n à l'échelle j

Introduction

La détection et la classification des événements dans un signal suscitent un intérêt grandissant depuis plusieurs années dans le domaine de l'ingénierie biomédicale. Pluridisciplinaire, elles progressent grâce aux avancées dans les divers domaines du traitement du signal. C'est dans le cadre des signaux biomédicaux que notre recherche s'inscrit : le but de ce travail est de développer un outil de recherche et de diagnostic fondé sur l'analyse des signaux électromyographiques (EMG) utérins. C'est un travail méthodologique de traitement de signal qui a comme objectif médical la prévention des accouchements prématurés. Ceux-ci représentent un souci majeur de la santé mondiale et nationale pour diminuer la mortalité et la morbidité (handicaps physiques et mentaux).

L'enregistrement de l'électromyogramme utérin permet de déceler l'activité musculaire de l'utérus gravide, susceptible de déclencher l'accouchement prématuré. Toutefois cet enregistrement représente la sommation électrique de toutes les activités contractiles et entre autres l'utérus et son contenu fœtal.

L'analyse d'enregistrements électriques est basée sur la détection et la classification des phénomènes transitoires et synchrones à des contractions plus au moins durables sur un fond d'activité électrique permanente d'origine non utérine.

La solution statistique d'un problème de détection dépend généralement de la connaissance *a priori* des paramètres de changement dans un signal donné. Quand les paramètres des distributions associées sont connus, des détecteurs optimaux basés sur le rapport de vraisemblance pourraient être utilisés. Dans le cas où ces paramètres seraient inconnus, des méthodes basées sur le rapport de vraisemblance généralisé consistant à utiliser une estimation du maximum de vraisemblance à partir de l'estimation des paramètres ont été utilisés dans la littérature. La modélisation du signal est donc une étape importante précédant la phase de détection. Elle permet d'explorer le signal pour en faire ressortir les paramètres de changement.

Les distributions fréquentielles et énergétiques sont les éléments essentiels qui permettent de détecter et identifier les événements dans le signal EMG utérin.

Une des modélisation faisant apparaître les changements fréquentiels et énergétiques est matérialisée par la décomposition en ondelettes et en paquets d'ondelettes.

L'application d'un algorithme de détection sur les paquets d'ondelettes permet en effet de détecter tout changement d'amplitude et de fréquence dans le signal, ce qui permet de classifier les événements détectés

Le fond du signal représente du bruit, et le choix d'une meilleure base à partir d'une décomposition en paquets d'ondelettes pour la détection permet de faciliter la détection des phénomènes transitoires spécifiques, d'où la nécessité d'élaborer un algorithme de sélection de la meilleure base.

Ayant défini les caractéristiques de la meilleure base adaptée à la détection, se pose alors la question du critère de sélection et de la stratégie de validation qui lui ont associés. La distance de Kullback Leibler nous a paru un critère pertinent pour la décision de présence d'événements dans un paquet donné puisqu'elle mesure la distance entre deux distributions statistiques.

Après cette sélection, une méthode de détection classique pourra être utilisée sur chaque élément de la base sélectionnée pour segmenter le signal et extraire les informations utiles. Une procédure de fusion sera cependant nécessaire pour synthétiser sur le signal global les détections obtenues sur chaque élément de la base.

Une fois les événements détectés, il est nécessaire de les classifier selon leur signification physiologique en utilisant un critère de classification permettant cette fois de choisir parmi les paquets d'ondelettes, les paquets les plus discriminants.

Après l'obtention des paquets discriminants, diverses méthodes de classification (K-plus proches voisins, distance de Mahalanobis, réseaux de neurones, et les "Support Vector Machines" (SVM)) sont utilisées pour faire correspondre chaque événement à la classe physiologique à laquelle il appartient.

En conséquence, ce mémoire est organisé de la manière suivante :

Le premier chapitre est une présentation globale du problème traité, qui peut s'énoncer comme étant la détection et la classification des événements dans le signal électromyographique (EMG) utérin. Ces événements aident en effet à la prévention de l'accouchement prématuré, ce qui est l'objectif final de ce travail.

Dans ce chapitre, nous commençons par la description de l'objectif global de l'étude consistant à prévenir les accouchements prématurés, puis un bref aperçu de la physiologie associée aux signaux EMG utérin est présenté. Ensuite la forme, le contenu, l'origine des signaux EMG et quelques définitions liées à la physiologie utérine utiles pour la suite du travail sont exposés.

La dernière partie de ce premier chapitre introduit les différentes étapes de traitement à mettre en œuvre pour la détection et la classification des événements de l'EMG utérin. Le but de cette analyse est discuté. Les différents choix adoptés pour réaliser le traitement de l'EMG sont justifiés.

Le deuxième chapitre constitue l'étape de modélisation du signal avant la détection et la classification. La première section de ce chapitre est dédiée à la décomposition discrète en ondelettes, consistant à décomposer le signal suivant des niveaux d'échelles sur lesquels se traduira tout changement de fréquence ou d'énergie. Le calcul des coefficients de détail et d'approximations est présenté brièvement.

La deuxième section aborde la décomposition en paquets d'ondelettes, intéressante dans la mesure où l'on peut ne conserver que les paquets qui reflètent bien le signal initial pour un objectif déterminé. Dans la décomposition en paquets d'ondelettes, chaque paquet représente l'information relative à tout le signal dans une bande de fréquence donnée. Le calcul des coefficients des paquets d'ondelettes est également rapidement présenté. Les principaux algorithmes du choix de la meilleure base, associés aux divers critères utilisés pour répondre à l'objectif de la sélection des meilleurs paquets, sont exposés. Cette présentation des algorithmes amène aux algorithmes du choix des meilleurs paquets pour la détection et la classification.

Dans le chapitre 3 une nouvelle approche de sélection de la meilleure base consistant à identifier les paquets révélant au mieux l'existence d'événements dans une classe donnée de signaux, est proposée. L'idée à la base de l'algorithme consiste à rechercher, parmi toutes les bases possibles, celle qui représente le banc de filtres le mieux adapté au signal en termes de mise en évidence des événements existant dans le signal de départ. Ayant défini les objectifs de détermination de la meilleure base capable de mettre en évidence des événements spécifiques dans un enregistrement, se pose alors la question du critère à utiliser et de la stratégie de sélection de la meilleure base qui lui est associée. Le critère utilisé sera la distance de Kullback Leibler (K-L) qui est un bon outil statistique pour décider de la capacité d'un

paquet à détecter les ruptures dans une classe de signaux donnée, compte tenu de sa capacité à mettre en évidence les différences de caractéristiques statistiques entre catégories d'événements.

Dans le chapitre 4, le principe retenu pour la détection est l'application d'algorithmes séquentiels sur les coefficients (et non sur les signaux reconstruits) des niveaux d'échelles (dans le cas de décomposition discrète en ondelettes) ou sur les meilleurs paquets préalablement choisis (dans le cas de décomposition en paquets d'ondelettes). La détection à partir des coefficients repose sur la recherche de la coïncidence en temps, à travers plusieurs échelles, de coefficients d'ondelettes significatifs. La décomposition dyadique peut cependant mener à un décalage important dans la localisation des changements. Le manque de synchronisme des détections au niveau de la séquence des coefficients des différents paquets ou niveaux d'échelles nous conduit à redéfinir la valeur vraie de l'instant de détection en référence au signal original.

Dans une première section de ce chapitre, la théorie du décalage des coefficients d'ondelettes sur chaque niveau d'échelle et sur chaque paquet d'ondelettes est étudiée. La correction des instants de détection sera ensuite abordée pour recalibrer les instants de changement à leur vraie position. Chaque instant de détection au niveau d'un paquet correspondra alors à un intervalle de détection au niveau du signal de départ. Comme la détection se fait sur plusieurs paquets ou niveaux d'échelles, un algorithme de fusion des intervalles de détection sera appliqué pour aboutir, lorsque cela sera justifié, à un seul instant de détection pour le signal original.

La faisabilité de la détection en temps réel sera ensuite démontrée. La détection sera traitée en temps réel. Enfin, la performance des méthodes est mesurée en les appliquant d'une part sur des signaux simulés, d'autre part sur la base de données disponible de signaux réels d'EMG utérin.

Enfin dans le chapitre 5, nous présentons le choix des paquets les plus discriminants obtenus à partir de l'ensemble des paquets d'ondelettes de l'arbre de décomposition. Le problème qui se pose est de définir un critère permettant le choix des paquets à forte valeur discriminante. Le rapport entre la variance intra classe et la variance totale est adopté comme un critère de discrimination. Il est appliqué sur l'ensemble d'apprentissage des événements EMG utérins.

Les approches de classification utilisant les K Plus Proches Voisins et la distance de Mahalanobis sont ensuite utilisées pour démontrer l'efficacité des paquets choisis à discriminer entre les différents événements. D'autres méthodes d'identification supervisée sont également présentées : les réseaux de neurones et la méthode SVM. Leurs principes sont

présentés, et les paramètres utilisés sont vérifiés. Des résultats sont obtenus et discutés en utilisant toute la base de données disponible. Nous terminons ce chapitre par une conclusion sur le critère de discrimination utilisé ainsi que les méthodes de classification appliquées sur les événements de l'EMG utérin en présentant leurs avantages et leurs limites.

Chapitre 1

Position du problème

1.1 Introduction

Ce premier chapitre est une présentation globale du problème traité, qui peut s'énoncer comme étant la détection et la classification des événements dans le signal électromyographique (EMG) utérin. Ces événements aident en effet à la prévention de l'accouchement prématuré, ce qui est l'objectif final de ce travail.

Dans ce chapitre, nous commençons par la description de l'objectif global de l'étude, puis nous donnerons un bref aperçu de la physiologie associée aux signaux EMG utérins. Ensuite nous présenterons la forme, le contenu et l'origine des signaux EMG et quelques définitions liées à la physiologie utérine utiles pour la suite du travail.

A la fin du chapitre, nous reviendrons sur les différentes étapes de traitement à mettre en œuvre pour la détection et la classification des événements de l'EMG utérin. Le but de cette analyse est discuté. Les différents choix adoptés pour réaliser le traitement de l'EMG sont justifiés.

1.2 Objectifs détaillés de l'étude

Cette étude s'inscrit dans le cadre de la prévention des menaces d'accouchements prématurés (MAP). La menace d'accouchement prématuré est un problème d'actualité car sur 750000 naissances en France, 44000 sont des naissances prématurées [Joh93].

D'après la définition de l'Organisation Mondiale de Santé (OMS), l'accouchement prématuré correspond à une durée de grossesse inférieure à 37 semaines complètes, ou inférieure à 259 jours d'aménorrhées. Tout accouchement survenant après 22 semaines d'Aménorrhée (SA) et avant 37 SA est ainsi défini comme un accouchement prématuré. Une naissance ayant lieu avant 22 semaines révolues d'aménorrhée est considérée comme un avortement pour l'OMS. La naissance prématurée est une pathologie pouvant occasionner de graves conséquences pour l'enfant mais aussi un coût socio-économique certain.

Les principaux risques pour l'enfant sont :

- Détresse respiratoire (souvent associée à une maladie des membranes hyalines)
- Infection
- Maladies neurologiques
- Hypothermie
- Entérocolite ulcéronécrosante

Depuis 20 ans, les progrès dans le suivi néonatal des enfants de moins de 1500 grammes ont favorisé considérablement leur taux de survie. L'un des problèmes majeurs auquel est confronté le monde obstétrical, dans le cadre de l'élaboration d'un traitement efficace, est que les causes de la MAP ne sont pas connues dans 40 % des cas [Ruf88]. Pourtant certains facteurs semblent jouer un rôle important dans la prévention des MAP.

Les facteurs socio-économiques, les infections microbiennes, les anomalies utérines, la rupture prématurée des membranes et les diverses pathologies de la grossesse sont des facteurs qui interviennent dans le risque d'accouchement prématuré.

La principale méthode de prévention de l'AP est fondée sur le dépistage et l'évaluation systématique des facteurs de risque de la femme. Ceux-ci sont évalués lors de la première consultation gynécologique de suivi de grossesse, par un interrogatoire de la femme sur son mode de vie associé à un examen clinique. La législation française oblige la femme à faire 7 consultations prénatales, ce qui permet au médecin d'avoir un suivi régulier de la grossesse et d'agir le plutôt possible d'une manière adaptée.

Actuellement, le suivi usuel de l'activité utérine au cours de la gestation est effectué par l'autopalpation de la femme et la tocographie externe

Cependant, l'autopalpation de la femme est de moins en moins utilisée du fait de son manque d'objectivité et de sensibilité. La tocographie externe, par nature objective, est fondée sur l'enregistrement des déformations de l'abdomen de la gestante au moyen d'un capteur de force, déformations provoquées par la contraction de l'utérus. Le problème majeur de cette

technique est qu'elle souffre d'un problème de sensibilité dans les termes très bas. De plus, elle est très sensible aux artefacts causés par les mouvements maternels et fœtaux. Ainsi, elle ne permet pas l'enregistrement des contractions de faible puissance. Ce type de contraction, rencontré au cours des termes précoces, conduit à une déformation de l'abdomen de la femme en dessous du seuil de sensibilité de l'appareil.

Toutes les femmes présentant un haut risque d'AP ou toute femme ressentant au moins deux contractions, parfois douloureuses, toutes les dix minutes pendant une heure, associées à une dilatation du col, sont soumises à un traitement tocolytique et hospitalisées si nécessaire. Le traitement tocolytique vise principalement à réduire le nombre de contractions. Son efficacité est fondée sur l'arrêt temporaire des contractions utérines, un retardement de l'accouchement de 24 à 48 heures, et si possible jusqu'à la 36^{ème} SA [Hig93].

Le monitoring de l'activité contractile utérine est un bon indicateur du risque d'AP [Hor96]. L'activité contractile du muscle utérin augmente physiologiquement au cours de la grossesse entre la 30^{ème} et la 44^{ème} SA [Nag88]. Nageotte a pu montrer, en 1988, que pour un âge gestationnel donné, l'activité utérine chez les femmes qui accoucheront prématurément est supérieure à celle qui accoucheront à terme ou au-delà du terme, indiquant que l'activité utérine des trois groupes est néanmoins identique à partir de trois jours avant l'accouchement. Cela montre donc bien que le suivi de l'activité utérine offre de bonnes perspectives pour la détection du risque d'accouchement prématuré. Cependant, il n'existe pas à l'heure actuelle de normes physiologiques du nombre de contractions par unité de temps.

Les méthodes d'évaluation et de suivi de la menace d'accouchement prématuré souffrent d'un grand problème de sensibilité et d'objectivité. De plus les méthodes actuelles de traitement ne sont réellement efficaces que lorsque la détection est précoce.

L'étude de la contractilité utérine par électromyographie (EMG) a été entamée depuis de nombreuses années. L'une des premières études date de 1931 [Bod31]. Les premières études se sont concentrées sur l'analyse temporelle puis fréquentielle de ce signal. Le signal électrohystérogaphique (EHG) a semblé être un vecteur potentiel d'indication du risque d'accouchement prématuré. En effet, il existe une bonne corrélation entre la survenue d'une bouffée d'activité électrique sur ce signal et l'augmentation de la pression intra-utérine [Ste50]. Les perspectives de prédiction de la menace d'accouchement prématuré avec ce signal ont néanmoins mis de nombreuses années d'études avant de se concrétiser. Cela a nécessité une longue phase d'analyse et d'élaboration de méthodes de traitement du signal EHG en vue de l'extraction de paramètres caractéristiques des contractions, et de la démonstration de leur efficacité.

En effet, le signal recueilli par électrodes de surface est soumis aux divers bruits liés à l'environnement (bruits électriques, mouvements d'électrodes) et contient des informations d'origines diverses desquelles il faut extraire uniquement les événements utiles au diagnostic. Par contre, par l'étude du contenu fréquentiel des divers événements recueillis par des électrodes de surface, il est possible de faire la distinction entre les contractions efficaces et inefficaces, et cela de manière non-invasive [Mar86]. De même cette analyse fréquentielle, associée à d'autres caractéristiques des événements (comme leur longueur) a déjà permis d'isoler les événements propres à la contraction utérine [Kha99]. L'idée est alors de chercher à améliorer les performances de détection et de classification actuelles en appliquant les méthodes de décomposition fréquentielle de ces signaux les plus efficaces, d'une part pour une détection optimale, d'autre part pour une sélection efficace des événements propres à la contraction utérine.

Le but de notre travail est donc de fournir des méthodes de détection et de classification qui permettent d'identifier les contractions parmi tous les événements existants dans le signal EMG utérin. Notons cependant que les autres ondes détectées par les électrodes ont également des influences perturbatrices. Certaines études se sont alors orientées vers la détection et la caractérisation de ces événements spécifiques [Mar86].

1.3 Description des signaux

L'étude de la contractilité utérine par électromyographie (EMG) a été entamée depuis de nombreuses années. Le travail de Bode [Bod31] en est une des premières études. Elles se sont concentrées sur l'analyse temporelle puis fréquentielle de ce signal, puisque le signal EMG a semblé être un vecteur potentiel d'indication du risque d'accouchement prématuré. En effet, il existe une bonne corrélation entre la survenue de bouffées d'activité électrique sur ce signal et l'augmentation de la pression intra-utérine [Ste50]. Il est donc important d'identifier les "bonnes" caractéristiques électriques des différents signaux composant les enregistrements afin de savoir les relier à un risque d'AP.

1.3.1 Signal EMG

Les premières études concernant le signal électromyographique utérin ont été réalisées par électromyographie interne. Elles ont permis la description du signal EMG, lors d'une contraction, comme une activité électrique dont la fréquence se situe entre 0.3 et 1 Hz et

d'amplitude comprise entre 100 μV et 1,8 mV [Lop84]. L'enregistrement simultané de l'activité interne et externe de l'EMG sur la femme a montré une très bonne corrélation entre les deux types de signaux [Hon58], [Wol70]. Cela a permis d'affirmer que le signal EMG de surface est représentatif de l'activité électrique du muscle utérin. Ainsi, l'EMG devient une nouvelle méthode d'étude, non-invasive, de l'activité électrique de l'utérus. Ces résultats ont été confirmés lors d'une analyse du signal EMG sur le macaque gravide [Man91].

L'EMG, enregistré à l'aide d'électrodes de surface, se caractérise par une activité basse fréquence (0,03 - 0,1 Hz), sur laquelle se superposent des séquences d'activité de plus haute fréquence (0,3 - 2Hz). Le signal de basse fréquence est considéré comme une activité résultante des perturbations mécaniques induites pour la déformation du ventre sous l'effet de contractions [Sur56]. Une étude sur la comparaison entre les contractions pendant la grossesse et lors de l'accouchement a montré que leur énergie se situe principalement dans la bande de fréquence 0,2 - 3Hz [Mar86]. Il a été montré qu'il existe néanmoins une différence de contenu fréquentiel entre les deux types de contraction (grossesse et accouchement). En effet, il y a un décalage vers les hautes fréquences du spectre de puissance des contractions, au fur et à mesure de l'avancement du terme de la femme [Mar86]. Ainsi, par l'étude du contenu fréquentiel des contractions, il est possible de faire la distinction entre les contractions efficaces et inefficaces de manière non-invasive pour des grossesses à terme. Par contre, l'analyse fréquentielle seule ne suffira pas toujours à distinguer les bouffées liées à la contraction utérine des autres événements contenus dans les enregistrements.

1.3.2 Les différents événements contenus dans l'EMG

Les enregistrements de l'EMG contiennent essentiellement deux types d'activités caractéristiques de la contraction utérine (bouffées de contractions et ondes d'Alvarez), mais aussi des artéfacts liés aux mouvements du fœtus. Il existe également des ondes dont l'origine n'est pas encore bien identifiée à ce jour, mais qu'il faut savoir isoler (LDBF). Une illustration de chaque type d'événement est représentée sur la figure 1.1.

Contractions

Ce sont des bouffées de longue durée (supérieure à une minute) et de fréquence d'apparition faible qui sont détectées sur les femmes dès la 30^{ème} semaine d'aménorrhée. Elles sont directement liées à la contraction utérine. Ces bouffées (notées par la suite de façon impropre

mais commode "contractions") présentent des caractéristiques spectrales de très basse fréquence (la quasi-totalité de l'énergie se situe à des fréquences inférieures à 0.5 Hz).

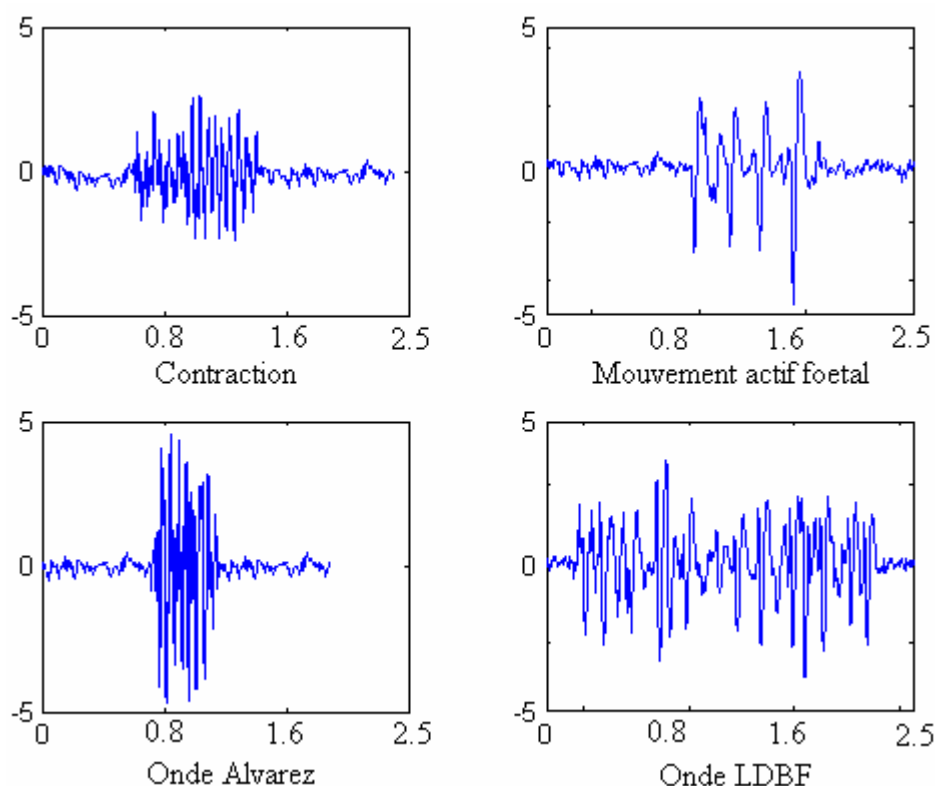


Figure 1.1. Les 4 types d'événements connus dans l'EMG utérin : une contraction, un mouvement actif foetal, une onde d'Alvarez, et une onde LDBF. Axes des abscisses : unités en minutes. Axes des ordonnées : unités arbitraires.

Les caractéristiques temporelles présentent aussi des différences en fonction du terme. Les bouffées d'activités associées aux contractions à mi-grossesse sont souvent de durée supérieure à 2 min, et elles présentent une répartition d'énergie dans une bande de fréquence inférieure à 0.2 Hz. Elles apparaissent comme une succession de bouffées plus courtes, étalées dans le temps, alors que les contractions des grossesses plus tardives sont associées à une seule bouffée ininterrompue de durée inférieure à 2 min qui présente une énergie significative au delà de 0.3 Hz (voir figure 1.2). Ceci peut être rapproché de l'évolution du mode de propagation de l'activité contractile utérine au cours du terme [Cal93].

Caractéristiques fréquentielles des contractions

Dans une bouffée relative à une contraction, on distingue généralement deux types d'ondes. La première (onde lente) principalement observée sur des enregistrements abdominaux, est

probablement due à des artefacts mécaniques. Elle couvre une bande de fréquences allant de 0.014 à 0.033 Hz. La seconde, appelée onde rapide, peut être enregistrée aussi bien en situation interne qu'externe, et pendant la grossesse comme pendant l'accouchement. Cette deuxième onde recouvre elle-même deux domaines spectraux différents correspondant, *a priori*, à deux événements physiologiques différents. Le domaine de plus basse fréquence, appelé FWL (pour "Fast Wave Low"), est présent dans tout le signal EMG. On tend à lier cette activité à des informations relatives à la propagation de l'activité électrique [Dev95a]. L'activité FWH (pour "Fast Wave High") est principalement observée dans des EMG correspondant à des contractions efficaces d'accouchement. Il semble à l'origine du contrôle de l'intensité de la contraction. Si l'on veut détecter des menaces d'accouchements prématurés et donc obtenir des informations liées à l'efficacité des contractions, il apparaît donc clairement nécessaire de s'intéresser à la présence de l'activité FWH. En effet, les caractéristiques de sa fréquence principale (énergie, forme de sa modulation) et surtout l'évolution de ses caractéristiques en fonction du contexte (espèce, grossesse/ accouchement, médications diverses) sont autant de paramètres riches de renseignements [Dev95b], [Duc95].

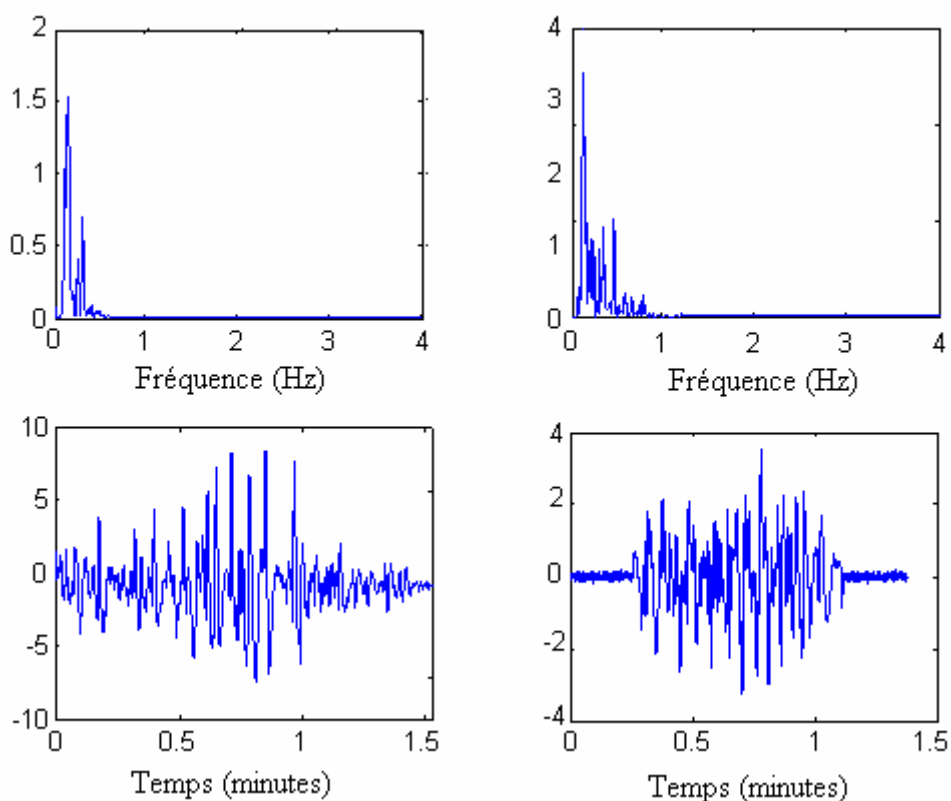


Figure 1.2. Gauche : contraction utérine à mi-grossesse et densité spectrale de puissance (DSP) associée. Droite : contraction utérine enregistrée à l'accouchement et densité spectrale associée. Axes des ordonnées : unités arbitraires.

Mouvements actifs fœtaux

Les bouffées correspondant aux mouvements actifs fœtaux (on notera également "MAF", toujours pour des raisons de commodité, les bouffées d'activité électrique correspondent à ces événements) ne constituent pas à proprement parler une composante intrinsèque du signal EMG de surface. Elles se traduisent cependant par des artefacts résultant de la déformation mécanique de l'abdomen de la femme sous l'action des mouvements du fœtus. Elles surviennent lors des phases d'activité naturelle du fœtus mais aussi comme réaction fœtale après la survenue des contractions sous l'effet de la contrainte mécanique générée. Il est à noter que les MAF peuvent provoquer la survenue de contractions si l'utérus de la femme est irritable [Mar95]. Les MAF ont un contenu spectral de très basse fréquence. De plus, leur caractéristique temporelle est très différente de celle des activités contractiles présentes dans l'EMG. Les MAF se présentent sous la forme d'un ou plusieurs pics de très courte durée et parfois de forte amplitude. La figure 1.3 présente deux types de mouvement actifs fœtaux avec leur densité spectrale associée. Le premier (à gauche) a déclenché des contractions, le deuxième (à droite) est provoqué par des contractions.

Ondes d'Alvarez

Les ondes d'Alvarez sont des bouffées d'activité de faible durée, allant de 30 secondes à 1 minute. Elles sont caractérisées par une haute fréquence, supérieure à 0.2 Hz. Ce type d'onde possède une fréquence d'apparition élevée (une toutes les une à deux minutes). Il s'agit d'une activité électrique locale, survenant de manière aléatoire en divers endroits de l'utérus. Il s'agit soit d'une activité physiologique normale, soit d'une hyperactivité pouvant aboutir à un accouchement prématuré [New87]. Les caractéristiques spectrales de ces ondes étant proches de celles des bouffées de contractions (FWH), il sera nécessaire de les différencier par d'autres caractéristiques. La durée de l'événement semble un bon critère dans ce cas. La figure 1.4 montre un exemple d'ondes d'Alvarez avec la densité spectrale associée.

Autres types d'activités

Les ondes LDBF. Ce sont des ondes qui ont une longue durée et un contenu de très basse fréquence en comparaison aux caractéristiques fréquentielles des bouffées liées aux

contractions. Jusqu'à maintenant la nature physiologique de ce type d'événement n'est pas documentée.

Les ondes Leman. Ce sont des ondes qui ont des caractéristiques fréquentielles proches de celles des contractions mais de faible amplitude. Elles seront généralement ignorées dans le processus de détection (rapport signal sur bruit très défavorable).

Au-delà des caractéristiques propres à chaque événement, il faut savoir tenir compte du fait que les caractéristiques des bouffées liées aux contractions varient d'une femme à l'autre et selon le terme de grossesse. De même, les mouvements actifs fœtaux et les bouffées liées aux contractions peuvent être confondus si on ne prend pas en compte le terme de grossesse. Enfin, il faut savoir prendre en compte le fait que la durée d'une contraction est inférieure à celle d'une onde LDBF et supérieure à une onde Alvarez.

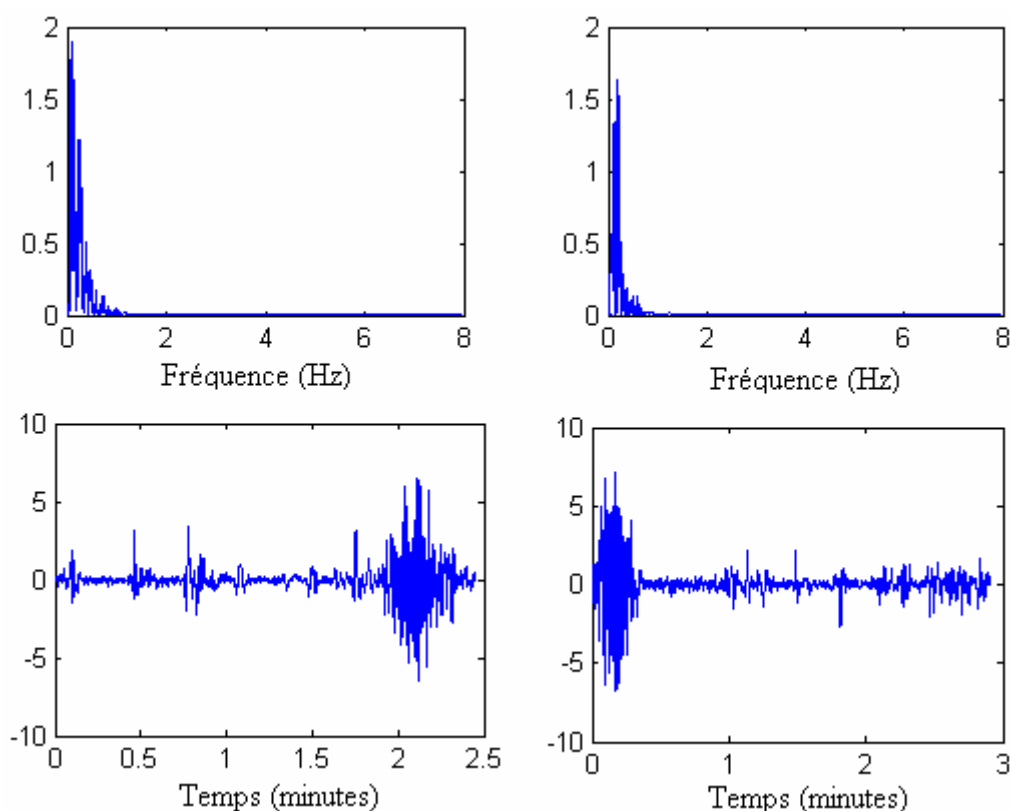


Figure 1.3. Deux exemples de mouvements actifs fœtaux. Gauche : MAF déclenchant une contraction avec la densité spectrale associée au MAF seulement. Droite : MAF provoqué par une contraction avec la densité spectrale associée au MAF seulement. Axes des ordonnées : unités arbitraires.

L'EMG semble donc un vecteur d'investigation prometteur dans le cadre de la détection de la menace d'accouchement prématuré. C'est ainsi que notre étude s'est concentrée sur la

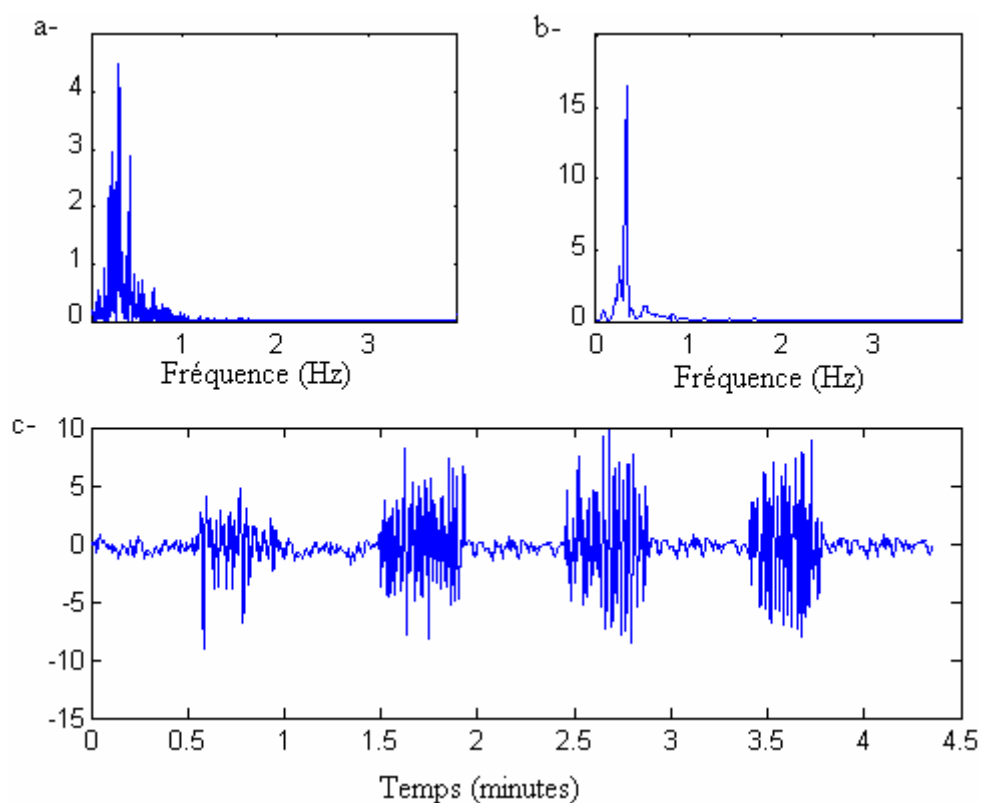


Figure 1.4. Ondes d'Alvarez identifiées dans un signal EMG utérin. (a) DSP associée à l'ensemble de la fenêtre temporelle représentée en (c). (b) DSP associée à la première bouffée de la fenêtre temporelle représentée en (c). Axes des ordonnées : unités arbitraire.

détection et la classification des événements de l'EMG pendant la grossesse, afin d'obtenir un moyen objectif de détection du risque d'accouchement prématuré.

1.4 Modélisation de l'EMG

Plusieurs modèles ont été développés pour la représentation des signaux EMG (et leurs méthodes d'estimation). La modélisation d'un signal EMG par un processus aléatoire est fréquemment utilisée par les chercheurs [Inb84], [Duc93], [Mer99].

Les objectifs essentiels de la modélisation d'un signal sont :

- la mise en évidence des caractéristiques énergétiques et fréquentielles du signal EMG.
- la description de son spectre par un ensemble très limité de paramètres [Kay88], [Mar87], [Orf90].
- l'utilisation des paramètres pour détecter des non stationnarités dans le signal [Obr88b] [Bas93], [Fal01], [Kha99b].
- l'utilisation des paramètres pour les problèmes de classification [Bas89].

Plusieurs méthodes existent dans la littérature pour le calcul des paramètres des différents modèles [Kay88], [Mar87], [Orf90]. Cependant, pour l'analyse du signal EMG, il a été montré qu'il était équivalent de modéliser ce signal par un modèle ARMA (Moving average autoregressive process) ou par un modèle Auto Régressif (AR) d'ordre plus élevé, puisque les coefficients du modèle autorégressif contiennent les informations sur le contenu fréquentiel et que la variance de l'erreur de prédiction permet de détecter tout changement d'énergie. Les méthodes de Yule-Walker, Burg, Covariance et Covariance modifiée sont capables de réaliser l'estimation des paramètres AR. Une comparaison de ces différentes méthodes d'estimation est disponible dans [Fal02].

Une deuxième approche de modélisation de l'EMG est fondée sur la décomposition en ondelettes. L'analyse par ondelettes est une méthode permettant d'analyser l'énergie de signaux non-stationnaires sur plusieurs plages de fréquence [Dau90]. Dans ce cas tout changement de fréquence ou d'énergie se traduira par un changement sur un ou plusieurs des niveaux de décomposition des signaux (niveaux d'échelles). Haar, dès 1910 [Mis04], a proposé la première décomposition en ondelette d'un signal (voir aussi [Mal99]). Actuellement, les ondelettes sont utilisées dans de nombreux domaines comme la compression de données [Hop94], le filtrage [Lem00],...

Une approche de modélisation à base d'ondelettes, mais produisant un arbre de décomposition, est la décomposition en paquets d'ondelettes. C'est cette approche qui sera testée au cours de ce travail, pour sa capacité à produire des sous arbres de décomposition performants pour la détection et la classification, tout en offrant des outils pour le prétraitement (débruitage des signaux).

Les paquets d'ondelettes permettent de décomposer le signal sur un grand nombre de bases et de choisir au sens d'un certain critère, celle qui représente au mieux le signal. Lors de la décomposition, les versions filtrées passe-bas et passe-haut du signal sont décomposées.

Les approximations des détails (filtrage passe bas) et les détails des détails (filtrage passe haut) viennent donc s'ajouter à l'approximation du signal. Les coefficients issus de cette décomposition sont caractérisés par trois paramètres : niveau de décomposition, indice de fréquence et indice temporel [Mis04]. Chaque paquet d'ondelettes est donc porteur d'une triple information $\{j, n, k\}$ échelle, fréquence, et position, là où l'ondelette ne comptait que deux paramètres, un paramètre d'échelle et un paramètre de position. Pour l'EMG utérin [Kha99] a utilisé la modélisation auto régressive et la décomposition en ondelettes alors que [Fal02] n'avait utilisé que la modélisation autorégressive. Dans le cadre de notre travail, la modélisation en paquets d'ondelettes a été utilisée pour tester si le choix d'une meilleure base

de paquets d'ondelettes pourrait améliorer les résultats en détection et classification. Une approche de sélection de meilleure base a donc été développée pour choisir les meilleurs paquets vis à vis des événements de l'EMG, d'une part en termes de détection, d'autre part en termes de classification.

1.5 Acquisition et Prétraitement de l'EMG utérin

Les différentes méthodes de traitement exposées dans les chapitres suivants seront testées sur des signaux réels acquis depuis plusieurs années par les équipes de l'université de technologie de Compiègne et du centre de gynécologie obstétrique de CHRU d'Amiens. Nous avons fait le choix de décrire le mode d'acquisition et de prétraitement des signaux dès ce chapitre, et de façon très succincte, car ces éléments ont été détaillés dans bon nombre de publications antérieures (voir par exemple [Kha99b], [Ter01]).

L'enregistrement des signaux EMG a été effectué au sein du Centre de Gynécologie Obstétrique de l'hôpital d'Amiens. Afin d'obtenir une population homogène et comparable, des critères d'exclusion et d'inclusion des patientes ont été clairement définis.

Les paramètres d'inclusion sont les femmes hospitalisées pour MAP et celles présentant des contractions. Les paramètres d'exclusion utilisés sont les femmes présentant une grossesse multiple et celles dont le terme est supérieur à 37 SA. Après avoir obtenu le consentement éclairé de chaque patiente, une préparation soigneuse de la peau a été effectuée au niveau de l'emplacement des électrodes. Les signaux ont été recueillis à l'aide d'une paire d'électrodes bipolaires apposées sur l'axe vertical médian de la paroi abdominale, en regard de l'utérus. Ces électrodes de surface étaient de type Ag/AgCl d'un diamètre de 8 mm. Les signaux étaient ensuite filtrés entre 0.2 et 6 Hz par un filtre analogique de façon à éliminer la composante continue et les artefacts dus au secteur puis amplifiés. Ils étaient ensuite numérisés avec une fréquence d'échantillonnage de 16 Hz [Kha99b].

Pendant la phase d'enregistrement, il était demandé à la femme de noter, à l'aide d'un marqueur d'événement manuel, tout mouvement de bébé ou toute contraction ressentie. Les marqueurs d'événements étaient enregistrés en parallèle du signal EMG dans une mémoire numérique.

Deux moniteurs d'enregistrement ont été utilisés :

- Le moniteur Myodata (Mazet électronique)
- Un moniteur expérimental (Hewlett-Packard HP)

Le moniteur HP a été développé dans le cadre de la prévention des MAP par électromyographie utérine [Voi00].

Les bases de différents signaux et événements utilisés sont décrits à l'annexe A.

1.6 Approches antérieures de traitement de l'EMG utérin

Une étape essentielle pour l'utilisation de l'EMG utérin dans le cadre de la Prévention des Menaces d'accouchement prématuré est donc de détecter et d'isoler les différents événements contenus dans le signal enregistré afin de les analyser ensuite de façon séparée. La phase de segmentation est donc cruciale dans le processus de traitement de ces signaux. Cette problématique est d'ailleurs générale et se retrouve dans de nombreux champs d'application.

Parmi les domaines où les chercheurs ont développé des algorithmes de segmentation, on peut citer le domaine de traitement de parole, le traitement des signaux sismiques et des signaux biomédicaux [Obr88a], [Obr88b], [Bas93], [Kha99a], [Kha99b].

Les algorithmes de détection sont essentiellement fondés sur la théorie statistique des tests d'hypothèses. Parmi ces algorithmes, nous nous sommes intéressés aux algorithmes séquentiels car ces algorithmes nous permettent de localiser les instants de changement même sur des signaux de longue durée. Citons par exemple l'algorithme de somme cumulée (CUSUM) qui est l'un des algorithmes séquentiels de détection les plus efficaces dans le cas où les paramètres sont connus [Bas93].

En pratique, les paramètres des distributions sous les différentes hypothèses ne sont pas connus et beaucoup d'algorithmes ont été développés pour résoudre ce type de problème [Bas93]. Khalil en 1999 [Kha99b] puis EL Falou en 2002 [Fal02], ont présenté une méthode originale permettant la détection et la classification de différents événements présents dans un signal non stationnaire. Cette méthode utilise deux techniques différentes de détection, l'une fondée sur une approche monodimensionnelle, l'autre sur une approche multidimensionnelle. Ce nouvel algorithme possède l'avantage de pouvoir détecter toute rupture fréquentielle et/ou d'énergie d'un signal et de classer ces différents événements de manière séquentielle, sans retourner sous l'hypothèse nulle H_0 (pas d'événement) pour toute nouvelle détection. Dans notre travail, nous cherchons à trouver d'abord une base qui représente l'EMG utérin selon ses caractéristiques en fréquence et en énergie, puis à appliquer sur cette base de décomposition l'algorithme de segmentation développé par [Kha99a] en nous appuyant sur les propriétés les plus pertinentes de l'EMG utérin.

1.7 Différentes étapes du traitement des signaux EMG

Comme nous l'avons déjà souligné, l'objectif global de l'étude est de développer une méthodologie pour la prévention des accouchements prématurés en s'appuyant sur les résultats de l'analyse de la contractilité utérine. Les activités utérines varient d'une femme à l'autre et en fonction du terme de grossesse. Par exemple, les mouvements actifs fœtaux n'ont de composantes fréquentielles différentes de celles des bouffées des contractions que pour un terme donné. Dépendant du terme, ces dernières ont en général des composantes fréquentielles plus élevées que celles des mouvements actifs fœtaux. La durée d'une contraction est supérieure à celle d'une onde d'Alvarez recueillie à un terme proche. Les contenus fréquentiels des MAF et des ondes LDBF sont très proches alors que la durée des MAF est inférieure à celle de LDBF. Il est donc important de considérer que les caractéristiques temporelles et fréquentielles changent en fonction du terme. La prévention des accouchements prématurés devra être abordée par la caractérisation des événements dans le signal EMG utérin à un terme donné. Pour arriver à ce but, il faut segmenter les enregistrements et isoler les événements utiles du signal global. Ceci implique la mise en œuvre d'une méthodologie de détection et de segmentation de ces événements, enregistrement par enregistrement, puis leur identification selon leur signification physiologique.

L'approche envisagée dans le cadre de ce travail est de décomposer d'abord l'enregistrement sur une base qui permette une détection efficace des événements en exploitant leur contenu fréquentiel, puis qui permette une classification de ces événements selon les catégories définies au §1.3.2. La décomposition en paquets d'ondelettes nous a semblé une piste très prometteuse dans ces objectifs à la fois de détection et de classification. Dans ce but, il est nécessaire de traiter les enregistrements en plusieurs étapes.

Les différentes étapes de traitement des enregistrements sont alors les suivantes.

Prétraitement

La première étape consiste à améliorer autant que possible le rapport signal sur bruit. Cette phase de "débruitage" est maintenant souvent abordée en s'appuyant sur une décomposition en ondelettes et en paquets d'ondelettes dans la littérature, et en particulier dans le cadre du traitement de l'EMG utérin [Lem97], [Lem98]. Le débruitage consiste à retenir un certain

nombre des coefficients et à éliminer les autres [Mal99]. Comme ces transformations en ondelettes et en paquets d'ondelettes offrent également un moyen simple de décomposition du signal en sous signaux de contenus fréquentiels différents, on peut envisager la même décomposition pour le débruitage et la segmentation. Les énergies sur les échelles ou sur les paquets correspondent au contenu fréquentiel, i.e. à l'énergie du signal original dans une bande de fréquence spécifique définie par chaque échelle ou chaque paquet. Utilisées dans un objectif de débruitage, ces décompositions peuvent donc servir également à la détection des événements.

Paquets d'ondelettes et détection

Abry et Flandrin [Abr94], [Abr97] se sont préoccupés de la détection en utilisant les ondelettes. Dans [Sen95b] une ondelette complexe a été utilisée pour la détection des ruptures de l'ECG. La segmentation est une étape très importante dans notre travail, l'objectif de la segmentation étant l'extraction des segments correspondant aux différents événements du signal EMG utérin. Compte tenu des spécificités fréquentielles de ces événements, il nous a paru opportun d'appliquer un algorithme de détection sur les coefficients d'ondelettes et les coefficients des paquets d'ondelettes, sans utiliser les signaux reconstruits. L'approche directe de détection sur les coefficients simplifie le calcul et permet une détection rapide des événements. Le principe de détection a déjà été appliqué sur des niveaux d'échelle sélectionnés dans le cas de décomposition en ondelettes [Kha99a]. Nous proposons une approche alternative pour laquelle la détection est effectuée sur chaque paquet d'une meilleure base dans le cas de décomposition en paquets d'ondelettes. La décomposition en paquets d'ondelettes est une approche plus fine de filtrage puisqu'elle permet de choisir les paquets qui offrent les caractéristiques fréquentielles et d'énergie de l'EMG utérin les plus pertinentes vis-à-vis de la détection, puis de la classification. La sélection des paquets pertinents sera effectuée en utilisant une base d'apprentissage.

Correction et fusion des instants de détection

Dans l'approche proposée dans [Kha99b] les signaux reconstruits ont été utilisés pour la détection, après décomposition en niveaux d'échelles. L'utilisation directe des coefficients pose un problème de recalage qu'il faut aborder avant la détection des instants de changement. Les coefficients sont en effet décalés sur chaque niveau d'échelle et sur chaque

paquet. Ce décalage est dû à plusieurs facteurs tels que la décimation en passant d'un niveau à un autre, le nombre de moments nuls des ondelettes utilisées et la forme de l'ondelette reliée à sa longueur et à sa symétrie. Une correction est alors nécessaire pour avoir des instants de détection qui correspondent à leur vraie place dans le signal original.

Dans la méthode abordée dans [Kha99b], la fusion était effectuée sur la fonction de détection multi échelle. Dans notre cas, compte tenu du fait qu'on a des niveaux indépendants, la matrice est diagonale, et de ce fait la détection peut s'appliquer en monodimensionnel sur chaque niveau. Par contre, la fusion prendra en compte la véritable résolution temporelle de chaque niveau due à l'utilisation des coefficients.

Par conséquent, un instant de détection à un niveau donné va correspondre à un intervalle de détection, rapporté au signal original. Après la correction des instants de détection sur les diverses échelles ou paquets d'ondelettes, un algorithme de fusion sera mis en œuvre pour obtenir un seul instant de détection pour l'ensemble des intervalles de détection identifiés sur les différents paquets ou niveaux d'échelles et correspondant, de fait, à l'identification du même événement.

Classification

Une fois les événements obtenus, la dernière étape est de les classifier selon leur signification physiologique. La classification revient à un problème classique, mais nous avons voulu bénéficier de la décomposition en paquets d'ondelettes pour sélectionner les paquets les plus discriminants basés sur un calcul de l'énergie. Chaque événement est alors décomposé lui-même sur un ensemble des paquets d'ondelettes. Sur la base d'un ensemble d'apprentissage, un critère de classification est appliqué pour obtenir l'ensemble des paquets les plus discriminants vis à vis des différents événements de l'EMG utérin. Après l'obtention de ces paquets discriminants, une méthode de classification classique (distance de Mahalanobis, réseaux de neurones...) est appliquée pour faire correspondre chaque événement à la classe physiologique (contraction, MAF, onde Alvarez, onde LDBF, bruit) à laquelle il appartient.

1.8 Synthèse

Ce chapitre a été consacré à une revue des données physiologiques nécessaires à la bonne compréhension des événements en jeu, et à une définition claire du problème et des étapes de

sa réalisation. Nous avons tout d'abord exprimé les objectifs généraux de l'étude, consistant à prévenir les accouchements prématurés.

Le signal EMG peut permettre un meilleur suivi de l'évolution des contractions au cours de la grossesse que celui apporté par les techniques actuellement utilisées (tocographie externe). En effet, comme nous l'avons vu précédemment, la détection précoce de la menace d'accouchement prématuré est primordiale pour la prévention.

Les méthodes de traitement de l'EMG utérin, consistant à détecter et à identifier les événements pertinents en utilisant les paquets d'ondelettes, ont été justifiées.

Dans le chapitre suivant, nous allons expliciter les méthodes de modélisation du signal utilisées pour mettre en évidence les composantes fréquentielles. La décomposition discrète en ondelettes est tout d'abord abordée, ensuite la décomposition en paquets d'ondelettes est présentée.

Chapitre 2

Représentations à base d'ondelettes et de paquets d'ondelettes

2.1 Introduction

Ce chapitre constitue l'étape de modélisation du signal avant la détection et la classification. Comme il est indiqué dans le chapitre précédent, les distributions de fréquence et d'énergie sont, avec la durée des événements, les éléments essentiels qui permettent de distinguer et d'identifier des événements dans les enregistrements de l'EMG utérin. Une modélisation est nécessaire pour mettre en évidence ces caractéristiques. Les détails obtenus après la décomposition en ondelettes, et les paquets obtenus après la décomposition en paquets d'ondelettes, permettent d'exploiter le contenu fréquentiel du signal. Dans ce cas en effet, tout changement de fréquence ou d'énergie se traduira par un changement sur un ou plusieurs niveaux d'échelles dans le cas de la décomposition discrète en ondelettes et sur un ou plusieurs paquets dans le cas de la décomposition en paquets d'ondelettes.

La décomposition en paquets d'ondelettes est intéressante dans la mesure où l'on peut ne conserver que les paquets qui reflètent bien le signal initial pour le critère choisi. Dans le cas de la détection, seuls les paquets contenant des événements sont retenus. Dans le cas de la classification, les paquets les plus discriminants sont choisis pour identifier les événements.

La première section de ce chapitre est dédiée à la décomposition discrète en ondelettes, consistant en la décomposition du signal suivant des niveaux d'échelles. Nous présentons brièvement l'analyse multirésolution, les filtres miroirs conjugués et les propriétés des

ondelettes orthogonales. Les algorithmes de calcul des coefficients des détails et des approximations sont mis en œuvre.

La deuxième section aborde la décomposition en paquets d'ondelettes. Le calcul des coefficients des paquets d'ondelettes est également rapidement présenté. Ensuite nous évoquons les algorithmes du choix de la meilleure base associés aux divers critères utilisés pour répondre à l'objectif de la sélection des meilleurs paquets.

2.2 Pourquoi les décompositions en ondelettes et en paquets d'ondelettes ?

Dans des domaines d'application très différents, la décomposition en ondelettes est utilisée à des fins de détection des transitoires [Abr97], [Lem95], [Kar95], [Lea95], [Tir97]. Elle a été utilisée dans les domaines de détection des contours dans les images [Mal92], dans la compression [Rio93] et le traitement de la parole [Kad92a], dans les applications biomédicales comme la détection du QRS dans l'ECG [Li95] ou l'analyse de l'EEG [Sen95a], [Car93], et de l'EMG [Lem97], [kha99a]. Les signaux électrocardiographiques (ECG) ont été largement analysés pour mettre en évidence les potentiels tardifs [Sir93], [Sen90], pour la détection des ondes QRS [Kad92b], et pour la réduction du bruit [Sen93]. Senhadji et al [Sen95b] ont reconstruit les modèles d'ECG à partir des points d'extrema des coefficients d'ondelettes. Gramatikov et *al.* [Gra95] ont appliqué la décomposition discrète en ondelettes (DWT) aux signaux ECG pour surveiller l'état des patients après une opération coronaire d'angioplastie. Pour chaque niveau d'échelle, les auteurs ont estimé l'énergie moyenne des signaux de détails, qui est normalisée par l'énergie moyenne du signal original pour un cycle cardiaque donné.

Kalayci and Ozdamar [Kal95] ont utilisé les signaux de détails de la transformation en ondelettes et les réseaux de neurones pour la détection des transitoires EEG.

[Tha93] a appliqué la décomposition en ondelettes pour examiner des latences aux réponses sensorielles, électriquement simulées, enregistrées sur le cortex cérébral chez le chat. Dans [Din01], le signal ECG est tout d'abord segmenté, puis la transformée discrète en ondelettes DWT est appliquée sur chaque segment [Kad99]. L'algorithme localise les maximas locaux de la valeur absolue de la DWT qui dépassent un seuil donné pour chaque niveau d'échelle. L'emploi des ondelettes a réduit d'une façon significative l'erreur dans la détection des complexes de QRS.

[Mer03] a effectué une représentation surfacique du signal EMG en utilisant la transformée continue d'ondelettes. En explicitant les propriétés des ondelettes, les régions au comportement lissé, modélisées par un polynôme de degré N , produiront des coefficients nuls ou faibles et rendront ces régions invisibles. Au contraire, les portions singulières du signal seront fortement marquées par la décomposition, et donc bien détectables. Des auteurs [Pri94], [Kri95] s'intéressent souvent à l'aspect détection dans un espace obtenu par transformation linéaire. D'autres analysent l'importance du pavage du plan temps fréquence dans les performances du détecteur [Boa90]. Abry et Flandrin [Abr94], [Abr97] ont constaté qu'une bonne coïncidence temporelle entre la position du changement et les coefficients d'ondelettes sur les différentes échelles est nécessaire pour avoir une détection optimale. Senhadji [Sen95a] a utilisé une ondelette complexe pour la détection des ruptures dans l'EEG. Ce choix s'est appuyé sur une expérimentation approfondie et fondée sur l'analyse de nombreuses portions du signal à étudier. La décomposition en ondelettes a été utilisée aussi pour le débruitage [Don95a], [Coif92], [Lem00] et comme un outil de classification [Sai94], [Sai95], [Sam95].

[Wei04] présente une étude sur la détection de décollement de composantes multicouches en utilisant une combinaison de la transformée en paquets d'ondelettes (WPT) et de l'analyse modale. [Rav01] utilise la segmentation par les ondelettes de Malvar pour la détection de transitoires réels marins. Dans [Rav98] Ravier développe un nouveau critère de recherche de base optimale pour décrire le signal en vue de la détection de transitoires. Il est fondé sur une estimation du cumulatif d'ordre 4. Cet outil conduit à imposer des contraintes de décorrélation et de gaussianité sur le bruit. Il utilise les paquets d'ondelettes pour segmenter les signaux selon leur représentation fréquentielle. Ceci permet de traiter le cas du bruit coloré et de se rapprocher ainsi du contexte expérimental en travaillant avec des enregistrements marins réels. [Lem00] utilise la WPT et propose des critères spécifiques pour le débruitage du signal EHG. Le critère proposé est basé sur le fait de minimiser l'énergie des coefficients du signal qui sont inférieurs à l'énergie des coefficients du bruit.

Malgré le fait qu'un grand nombre de chercheurs utilisent les signaux reconstruits, quelques uns par contre utilisent directement les coefficients d'ondelettes. [Mol02], [Dau01] ont profité des propriétés spécifiques des coefficients d'ondelettes et ont proposé un cadre général pour des modèles de transitoires, fondés sur des arbres dyadiques de coefficients d'ondelettes. Les coefficients sont modélisés comme des variables aléatoires suivant un mélange de lois normales. [Let00] a présenté une nouvelle méthode appelée "Wavelet based Spike Classifier" (WSC) utilisant les coefficients d'ondelettes comme une base d'apprentissage pour classifier

les transitoires des signaux neuronaux "neural signals" dans l'espace temps fréquence. [Alb03] présente un codeur d'image localement adaptatif, basé sur la quantification des coefficients d'ondelettes. [Ran01], [Ran03] a étudié le développement d'un système d'étude de sons. Les pré-traitements tels que la détection, la segmentation et le débruitage par ondelettes ont été réalisés par une étude directe sur un vecteur de coefficients d'ondelettes.

2.3 Les bases de la décomposition discrète en ondelettes

La transformée en ondelettes permet d'analyser efficacement des signaux où se combinent des phénomènes d'échelles très différentes. La transformée en ondelettes continue d'un signal $f(t)$ prend la forme [Tru98] :

$$T^{ond} f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt \quad (2.1)$$

Dans cette expression, a est le facteur d'échelle et b le paramètre de translation. La variable a joue le rôle de l'inverse de la fréquence. L'expression (2.1) peut être également interprétée comme une projection du signal sur une famille de fonctions analysantes $\psi_{a,b}$ construites à partir d'une fonction "mère" ψ conformément à l'équation suivante :

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2.2)$$

Pour une transformation discrète en ondelettes, posons :

- $a = a_0^j$ avec $a_0 > 1$
- $b = k.b_0.a_0^j$ avec $b_0 > 0$, avec j et k entiers.

Dans la suite du travail, nous prenons $a_0 = 2$ et $b_0 = 1$ pour une transformation dyadique.

Le principe de la décomposition discrète en ondelettes est alors de construire des ondelettes ψ qui génèrent des bases orthonormées de $L^2(\mathbb{R})$ par translation et par dilatation [Mal99]:

$$\left\{ \psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t-2^j k}{2^j}\right) \right\}_{(j,k) \in \mathbb{Z}^2} \quad (2.3)$$

Les ondelettes orthogonales dilatées de 2^j reproduisent les variations d'un signal à la résolution 2^{-j} . La construction de ces bases est ainsi liée à l'analyse multirésolution des signaux. Ce lien nous conduit à une équivalence entre les bases d'ondelettes et les filtres miroirs conjugués utilisés dans les bancs de filtres sous-échantillonnés. Ces bancs de filtres implantent une transformée rapide en ondelettes orthogonales, qui ne nécessite que $O(M)$ calculs pour un signal de taille M . La construction de filtres miroirs conjugués permet de trouver de nouvelles bases orthogonales d'ondelettes, notamment avec des ondelettes régulières à support compact.

Pour $f \in L^2(\mathbb{R})$, la somme $\sum_{k=-\infty}^{+\infty} \langle f, \psi_{j,k} \rangle \psi_{j,k}$ de coefficients d'ondelettes peut s'interpréter comme la différence entre deux approximations aux résolutions 2^{-j+1} et 2^{-j} [Mal99]. Les approximations multirésolution calculent l'approximation d'un signal à diverses résolutions par projection orthogonale sur une famille d'espaces $\{V_j\}_{j \in \mathbb{Z}}$. Les approximations multirésolution sont entièrement caractérisées par un filtre discret particulier qui détermine la perte d'information entre deux résolutions. Ces filtres fournissent une procédure simple pour construire des bases orthogonales d'ondelettes.

La fonction d'échelle ϕ est définie par:

$$\phi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \phi\left(\frac{t-k}{2^j}\right) \quad (2.4)$$

La famille $\{\phi_{j,k}\}_{k \in \mathbb{Z}}$ est alors une base orthonormée de V_j pour tout $j \in \mathbb{Z}$.

2.3.1 Analyse multirésolution

L'analyse multirésolution est construite à l'aide de sous espaces V_j emboîtés les uns dans les autres, de manière que le passage de l'un à l'autre soit le résultat d'un changement d'échelle (zoom). Une suite $\{V_j\}_{j \in \mathbb{Z}}$ de sous espaces fermés $L^2(\mathfrak{R})$ est une approximation multirésolution si elle vérifie les propriétés suivantes [Mal99] :

- $V_{j+1} \subset V_j$: les V_j sont emboîtés les uns dans les autres.
- $\bigcap_j V_j = \{0\}$: le seul élément commun à tous les V_j est la fonction nulle partout.

- $\cup_j V_j = L^2(\mathfrak{R})$: pour toute fonction de $L^2(\mathfrak{R})$ on peut trouver au moins un des ensembles V_j la contenant.
- $f(\cdot) \in V_j \Leftrightarrow f(2\cdot) \in V_{j+1}$: si la fonction f appartient à V_j , la même fonction dilatée d'un rapport 2 appartient à V_{j+1} .
- $f(\cdot) \in V_0 \Leftrightarrow f(\cdot - n) \in V_0$: si la fonction f appartient à V_0 alors la même fonction traduite de n est encore dans V_0 .
- de plus il doit exister une fonction $\phi(t)$ telle que l'ensemble des fonctions $\{\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k)\}$ constitue une base orthonormée de V_0 . Cette fonction ϕ est appelée fonction échelle.

2.3.2 Filtrés miroirs conjugués

Toute fonction d'échelle est déterminée par un filtre discret appelé filtre miroir conjugué. La propriété de causalité des analyses multirésolutions impose que $V_j \subset V_{j-1}$. En particulier, $2^{-1/2} \phi(t/2) \in V_1 \subset V_0$. Comme $\{\phi(t-k)\}_{k \in \mathbb{Z}}$ est une base orthonormée de V_0 , on peut faire la décomposition:

$$\frac{1}{\sqrt{2}} \phi\left(\frac{t}{2}\right) = \sum_{k=-\infty}^{+\infty} h(k) \phi(t-k), \quad (2.5)$$

avec

$$h(k) = \left\langle \frac{1}{\sqrt{2}} \phi\left(\frac{t}{2}\right), \phi(t-k) \right\rangle. \quad (2.6)$$

Cette équation d'échelle exprime la dilatation de ϕ par 2 en fonction de ses translations entières. On interprète la suite $h(k)$ comme un filtre discret.

Les approximations de f aux échelles 2^j et 2^{j-1} sont données respectivement par la projection de f sur V_j et sur V_{j-1} . Soit W_j le complémentaire orthogonal de V_j dans V_{j-1} :

$$V_{j-1} = V_j \oplus W_j \quad (2.7)$$

Le théorème de Mallat [Mal89], [Mey90] montre que l'on peut construire une base orthonormée de W_j par translation et dilatation d'une ondelette ψ .

Théorème (Mallat, Meyer)

Soit ϕ une fonction d'échelle et h son filtre miroir conjugué. Soit ψ la fonction dont la transformée de Fourier vaut :

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}} \hat{g}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right) \quad (2.8)$$

avec

$$\hat{g}(\omega) = e^{-i\omega} \hat{h}^*(\omega + \pi) \quad (2.9)$$

Pour toute échelle 2^j , la famille $\{\psi_{j,k}\}_{k \in \mathbb{Z}}$ définie à l'équation (2.3) forme une base orthonormée de W_j . Sur l'ensemble des échelles, $\{\psi_{j,k}\}_{(j,k) \in \mathbb{Z}^2}$ forme une base orthonormée de $L^2(\mathbb{R})$.

La preuve du théorème montre que \hat{g} est la série de Fourier de

$$g(k) = \left\langle \frac{1}{\sqrt{2}} \psi\left(\frac{t}{2}\right), \phi(t-k) \right\rangle \quad (2.10)$$

qui sont les coefficients de la décomposition suivante :

$$\frac{1}{\sqrt{2}} \psi\left(\frac{t}{2}\right) = \sum_{k=-\infty}^{+\infty} g[k] \phi(t-k). \quad (2.11)$$

En prenant la transformée de Fourier inverse de (2.9) on obtient

$$g(k) = (-1)^{l-k} h(l-k). \quad (2.12)$$

Ce filtre miroir joue un rôle important dans les algorithmes de transformées en ondelettes rapides.

2.3.3 Propriétés et caractéristiques des ondelettes orthogonales

Les ondelettes ont quelques caractéristiques telles que la localisation, compacité du support, ..., etc. Nous en décrivons les plus importantes dans ce qui suit.

Localisation: L'analyse par ondelettes consiste en une décomposition/reconstruction du signal à l'aide de briques élémentaires d'encombrement limité simultanément dans les domaines temporel et spectral. La qualité de cette localisation conjointe est mesurée par rapport à la borne inférieure du principe d'incertitude de Gabor-Heisenberg.

Compacité du support: Il est souhaitable que l'ondelette ait un support fini, dans le domaine temporel ou fréquentiel. Cela permet de garantir efficacité et précision dans le calcul effectif des coefficients d'ondelettes et limite les problèmes de troncature.

Régularité: Dans les deux cas de décomposition et de reconstruction en utilisant les ondelettes, il est naturel de souhaiter que l'ondelette présente certaines qualités de régularité. Il est préférable, par exemple, que les r premières dérivées de l'ondelette soient bornées :

$$\forall n = 0, \dots, r, \left| \frac{d^n}{dt^n} \psi_0(t) \right| \leq C_n \quad (2.13)$$

Notons que la régularité temporelle de l'ondelette est liée à sa localisation fréquentielle. L'inverse est également valide.

Nombre de moments nuls: le nombre de moments nuls de l'ondelette est l'entier N tel que:

$$\forall k \in \{0, \dots, N-1\}, \int t^k \psi(t) dt = 0 \quad (2.14)$$

Ce nombre N implique que la transformée en ondelettes est aveugle aux portions douces d'un signal et que les coefficients n'en capturent que les comportements les plus singuliers.

Symétrie: Il est souhaitable parfois que les ondelettes présentent des propriétés de symétrie temporelle afin qu'elles constituent des bancs de filtres n'induisant pas de distorsion de phase. Notons que dans le cas d'ondelettes issues de multirésolution, parité ou imparité signifie

symétrie par rapport à $t=1/2$ et non $t=0$. L'origine de cette particularité est liée au schéma de décomposition/reconstruction exacte [Abr97]. Nous verrons par la suite que, pour les objectifs de détection, il sera possible de corriger cette distorsion de phase introduite par des filtres dissymétriques.

Orthogonalité: Cette propriété est à l'origine de la popularité des ondelettes qui offrent des bases non redondantes et orthogonales d'atomes temps-fréquence et donc réussissent là où le temps fréquence peut échouer.

2.3.4 Décomposition en ondelettes orthogonales : Algorithme pyramidal de Mallat

2.3.4.1 Définition

Pour effectuer concrètement la décomposition, on ne dispose pas en général du signal mais seulement d'une approximation à une échelle donnée.

Il faut donc trouver un algorithme qui, à partir de cette approximation, permet de trouver l'approximation et les détails à la résolution immédiatement inférieure, ce qui par itérations successives conduit à l'analyse du signal pour toutes les résolutions inférieures à celle de départ. Deux algorithmes principaux ont été mis en évidence : l'algorithme à trous [She92] et l'algorithme de Mallat [Mal89a], [Mal89b], [Mal99]. Le premier concerne des analyses multirésolution non orthogonales, le second est pratiquement le seul utilisé dans le cas des analyses orthogonales et biorthogonales. Nous présentons dans le paragraphe suivant l'algorithme de Mallat.

2.3.4.2 Calcul des coefficients de la transformée discrète en ondelettes : projection sur les fonctions d'échelle

L'action du projecteur sur une fonction f fournira la décomposition sur la base des fonctions d'échelles et les coefficients de cette décomposition constituent l'approximation à l'échelle j de f .

$$A_j f = \sum_k \langle f, \phi_{j,k} \rangle \phi_{j,k} \quad (2.15)$$

La formule de calcul des coefficients des approximations peut s'écrire

$$a_j(k) = \langle f, \phi_{j,k} \rangle \quad (2.16)$$

L'approximation à la résolution j de la fonction f sera définie par la suite discrète des nombres (réels ou complexes) $a_j(k)$.

Le point clef est fourni par la décomposition de $a_j(k) = \langle f, \phi_{j,k} \rangle$ en fonction de $a_{j-1}(k) = \langle f, \phi_{j-1,k} \rangle$.

Par construction, $\phi(x)$ est une fonction de V_0 . Comme $V_0 \subset V_{-1}$ on peut décomposer $\phi(x)$ sur la base de V_{-1} , et donc $\exists h(k)$ suite numérique avec $k \in \mathbb{Z}$ telle que:

$$\phi(x) = \sum_k h(k) \phi_{-1,k}(x) \quad (2.17)$$

mais comme $\phi_{-1,k}(x) = 2^{-1/2} \phi(2x - k)$ alors $h(k) = \langle \phi, \phi_{-1,k} \rangle$.

La suite numérique $h(k)$ sera considérée comme étant la réponse impulsionnelle d'un filtre numérique.

Montrons que la décomposition est la même pour des échelles quelconques.

On a :

$$\phi(x) = \sum_k h(k) 2^{1/2} \phi(2x - k) \quad (2.18)$$

donc :

$$\phi_{j,k}(x) = 2^{-j/2} \sum_n h(n) 2^{1/2} \phi(2(2^{-j}x - k) - n) \quad (2.19)$$

Ce qui, en regroupant les indices et les exposants, conduit à l'équation :

$$\phi_{j,k} = \sum_n h(n) \phi_{j-1, n+2k} \quad (2.20)$$

Les coefficients $a_j(k) = \langle f, \phi_{j,k} \rangle$ de l'approximation à la résolution j , sont donc calculés comme:

$$a_j(k) = \sum_n h(n) \langle f, \phi_{j-1, n+2k} \rangle \quad (2.21)$$

Si on pose $l = n+2k$, cette expression s'écrit :

$$a_j(k) = \sum_l h(l-2k) \langle f, \phi_{j-1, l} \rangle \quad (2.22)$$

Et si on note $\hat{h}[k] = h[-k]$ la séquence retournée ou le filtre symétrique de h , on obtient :

$$a_j(k) = \sum_l \hat{h}(2k-l) \langle f, \phi_{j-1, l} \rangle \quad (2.23)$$

et on aura finalement l'équation récursive suivante :

$$a_j(k) = \sum_l \hat{h}(2k-l) a_{j-1}(l) \quad (2.24)$$

Si on considère $a_j(k)$ comme une séquence numérique indexée par k , le calcul peut être interprété comme un produit de convolution entre \hat{h} et a_{j-1} évalué pour un indice sur deux, ou encore comme le filtrage de la séquence a_{j-1} par le filtre de réponse impulsionnelle \hat{h} suivi par un sous-échantillonnage de rapport 2.

2.3.4.3 Calcul des coefficients de la transformée discrète en ondelettes : Projection sur les fonctions d'ondelettes

Un schéma analogue est bâti à partir de la décomposition de l'ondelette de W_0 sur la base de V_{-1} :

$$\psi = \sum_k g(k) \phi_{-1, k} \quad (2.25)$$

ou de façon plus détaillée :

$$\psi(x) = \sum_k g(k) \sqrt{2} \phi(2x - k) \quad (2.26)$$

Ce qui conduit à l'équation suivante:

$$g(k) = \langle \psi, \phi_{-l,k} \rangle \quad (2.27)$$

$g(k)$ sera également considérée comme la réponse impulsionnelle d'un filtre numérique dont on peut caractériser l'ondelette $\psi(t)$ par ses coefficients. Ce filtre est lié au filtre $h(k)$ et il peut être reconstruit à partir de ce dernier comme l'on a déjà vu plus haut.

A noter que la fonction d'échelle $\phi(t)$ et son filtre associé $h(k)$ possèdent les caractéristiques d'un filtre passe-bas tandis que l'ondelette $\psi(t)$ et son filtre associé $g(k)$ possèdent celles d'un filtre passe-haut.

Un calcul analogue en tous points au précédent permet d'écrire les coefficients de détail :

$$d_j(k) = \langle f, \psi_{j,k} \rangle \quad (2.28)$$

$$d_j(k) = \sum_n g(n) \langle f, \phi_{j-l,n+2k} \rangle \quad (2.29)$$

On introduit également le filtre symétrique dont la réponse impulsionnelle correspond à la séquence $g(k)$ retournée :

$$\hat{g}(k) = g(-k) \quad (2.30)$$

La décomposition en ondelettes à l'échelle j pourra s'écrire:

$$d_j(k) = \sum_l \hat{g}(2k-l) \langle f, \phi_{j-l,l} \rangle \quad (2.31)$$

ou encore :

$$d_j(k) = \sum_l \hat{g}(2k-l) a_{j-l}(l) \quad (2.32)$$

Cette relation sera interprétée de la même manière que précédemment.

La figure 2.1 montre la décomposition d'un signal synthétique formé de trois segments de bruit blanc gaussien de même variance mais filtrés dans des bandes de fréquences différentes. Cette décomposition est présentée sur quatre niveaux de détails et un niveau d'approximation.

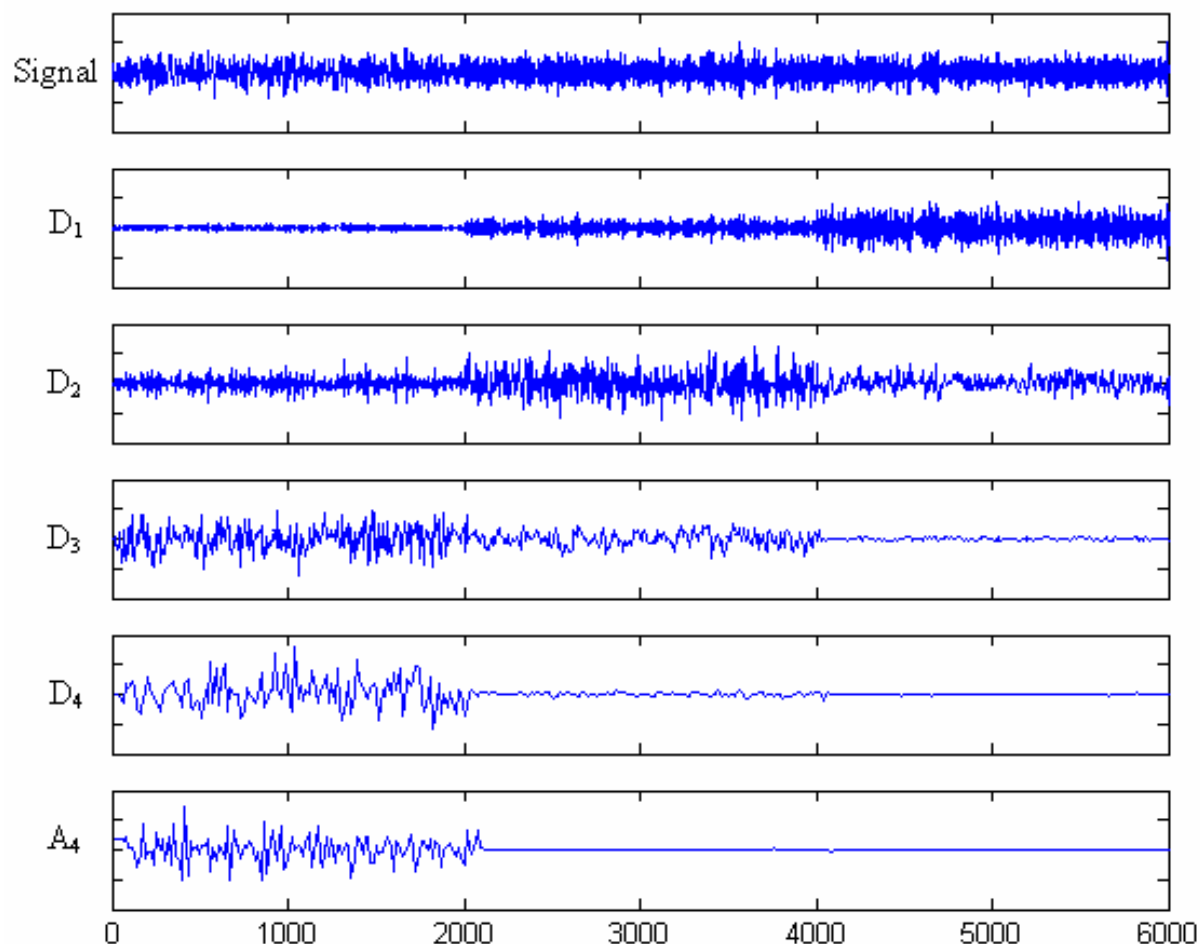


Figure 2.1. Décomposition d'un signal synthétique en quatre niveaux de détails et un niveau d'approximation. Axes des abscisses : nombre des points. Axes des ordonnées : unités arbitraires.

D_1 , D_2 , D_3 et D_4 sont les coefficients de détails, A_4 contient les coefficients d'approximation au niveau 4. Un changement fréquentiel apparaît comme changement d'énergie sur les échelles. La longueur du signal est 6000 points, les longueurs respectives de D_1 , D_2 , D_3 et D_4 sont de 3000, 1500, 750 et 375 points respectivement. La longueur de A_4 est de 375 points. Les coefficients sont placés en référence au signal original. A chaque 2^i points du signal est placé un point de D_i . De même pour A_4 : à chaque 16 points du signal est placé un point de A_4 .

2.4 Les paquets d'ondelettes

La caractéristique principale de la transformée en ondelettes est de permettre une analyse à des résolutions variables selon les échelles d'analyse. Un phénomène lent de longue durée ne nécessite pas une bonne localisation temporelle, tandis que les événements rapides sont traités avec une résolution temporelle précise, au détriment de la précision fréquentielle. L'approche par paquets d'ondelettes est différente : elle offre une librairie d'ondelettes organisées en fonction de leurs propriétés d'analyse et de localisation temps-fréquence, et donc de filtrage passe-bande, selon une architecture d'arbre binaire. Cette architecture permet alors de mettre en œuvre des algorithmes de recherche de bases adaptées à la fois aux propriétés temps-fréquence souhaitées et au signal analysé, que l'on dénomme classiquement "meilleures bases". Le dépassement de la structure dyadique nécessite l'introduction de nouvelles fonctions de bases, des ondelettes issues de l'ondelette mère, non plus simplement par dilatation, mais également par des opérations de convolution. Le problème d'avoir une segmentation fréquentielle dynamiquement adaptée au signal, est donc dû à la dualité de la problématique d'adaptation d'un banc de filtre à un signal : il faut d'une part un critère permettant de choisir la base fréquentielle la mieux adaptée, d'autre part, une méthode de construction de la base retenue.

De plus, cette décomposition présente l'avantage d'être définie et mise en œuvre à partir d'outils théoriques et d'algorithmes développés pour la transformée discrète en ondelettes (DWT).

2.4.1 Décomposition en paquets d'ondelettes

Les paquets d'ondelettes ont été définis par Coifman, Meyer et Wickerhauser [Mal99] en généralisant la relation entre les approximations multirésolution et les ondelettes. Un espace V_j d'une approximation multirésolution se décompose en un espace de résolution plus basse V_{j+1} et un espace de détails W_{j+1} . Cela se fait en divisant la base orthogonale $\{\phi_j(t - 2^j k)\}_{k \in \mathbb{Z}}$ de V_j en deux nouvelles bases orthogonales $\{\phi_{j+1}(t - 2^{j+1} k)\}_{k \in \mathbb{Z}}$ de V_{j+1} et $\{\psi_{j+1}(t - 2^{j+1} k)\}_{k \in \mathbb{Z}}$ de W_{j+1} . Alors que la décomposition en ondelettes se contente de décomposer les espaces d'approximations V_j et garde intacts les espaces de détails W_j , la

décomposition en paquets d'ondelettes opère encore sur les espaces de détails pour en déduire de nouvelles bases.

Les figures 2.2 et 2.3 décrivent d'une manière simple la différence entre la décomposition en ondelettes et celle en paquets d'ondelettes. La première consiste à décomposer V_j en V_{j+1} et W_{j+1} puis à réitérer cette procédure uniquement pour V_{j+1} puis V_{j+2} jusqu'à l'obtention de V_J et W_J alors que la deuxième applique cette décomposition à V_{j+1} mais aussi à W_{j+1} et redécompose chacun des sous espaces obtenus en utilisant le même banc de filtres associé à l'analyse multirésolution.

Le partage récursif des espaces vectoriels dans le cas de décomposition discrète en ondelettes se représente par un arbre binaire comme indiqué à la figure 2.2.

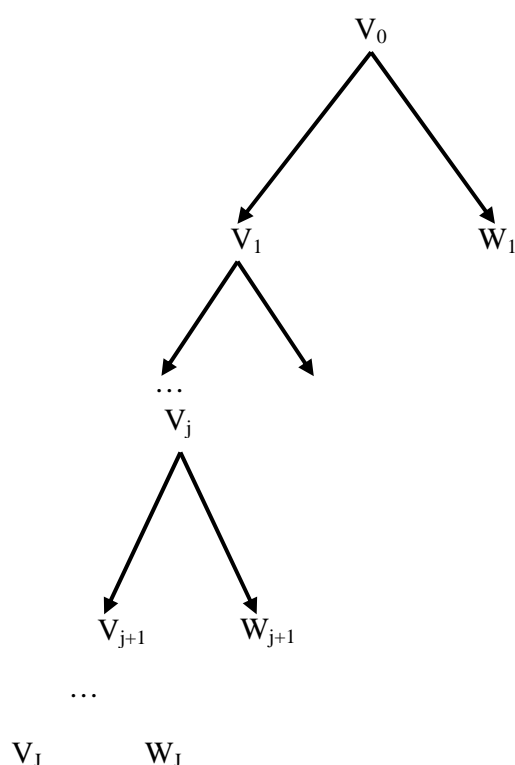


Figure 2.2. Décomposition en ondelettes à l'ordre J .

Pour mieux illustrer la décomposition en paquets d'ondelettes associée à la construction des bases, on donne le schéma suivant :

Soit $j \in \mathbb{Z}$, pour chaque échelle j , les valeurs possibles du paramètre n qui numérote les paquets sont : $0, 1, \dots, 2^j - 1$. Dans la suite dans le cas de décomposition en paquets d'ondelettes, l'espace V sera remplacé par W (figure 2.3).

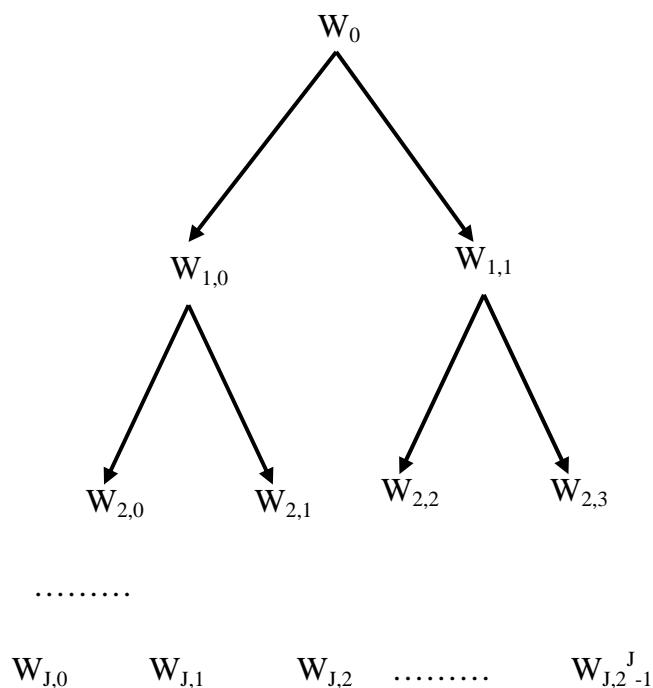


Figure 2.3. Décomposition en paquets d'ondelettes à l'ordre J .

Les paquets d'ondelettes sont indexés par trois paramètres : l'échelle, l'index et la position. j peut être interprété comme un paramètre d'échelle, n comme un index de chaque nœud de l'échelle j et k comme un paramètre de localisation temporelle. Pour les différentes valeurs de j et n la décomposition en paquets d'ondelettes est organisée en arbres. L'arbre tracé à la figure 2.4 est créé pour un niveau de décomposition égal à quatre.

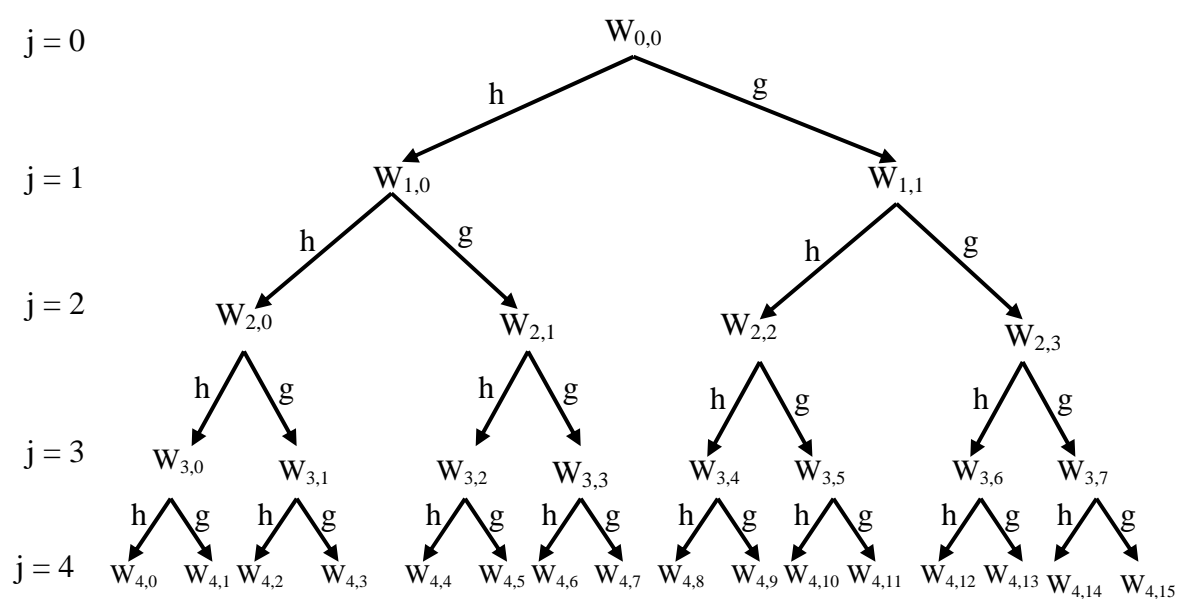


Figure 2.4. Arbre obtenu après décomposition en paquets d'ondelettes.

L'arbre obtenu après décomposition en paquets d'ondelettes peut se présenter en introduisant le niveau d'échelle et le numéro de paquet pour chaque niveau de décomposition, c'est le cas de la figure 2.4. Une autre façon de représenter l'arbre est de numéroter les paquets par ordre croissant (figure 2.5).

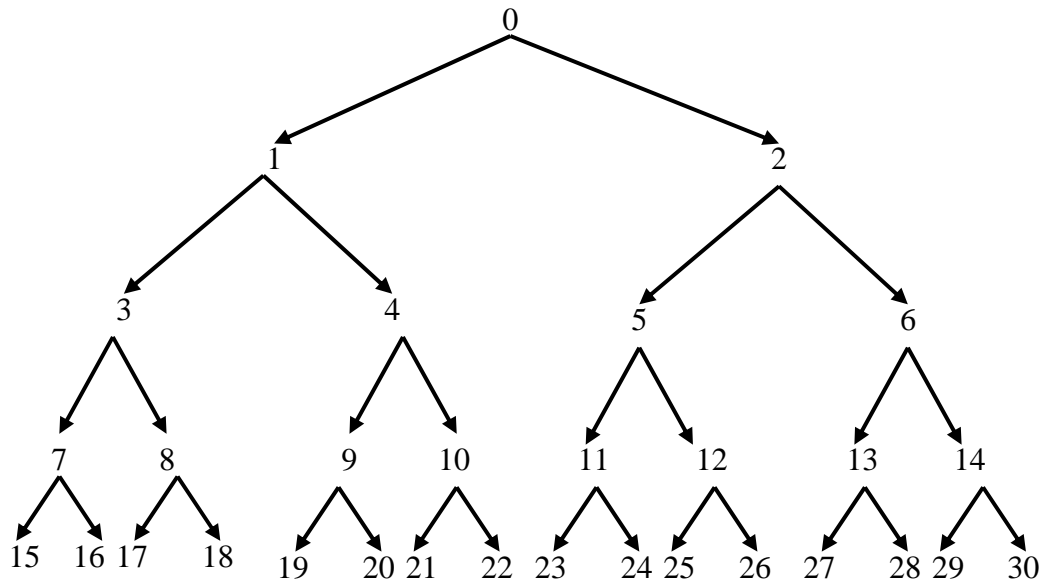


Figure 2.5. Arbre obtenu après décomposition en paquets d'ondelettes, paquets numérotés.

2.4.2 Bancs de filtres de paquets d'ondelettes

Chaque nœud d'un arbre binaire est repéré par les indices (j, n) où n est le numéro du nœud à la profondeur j . A chaque nœud est associé un espace $W_{j,n}$ qui admet une base orthonormale $\{\psi_{j,n}(t - 2^j k)\}_{k \in \mathbb{Z}}$. Les deux bases orthonormées de paquets d'ondelettes aux nœuds fils du nœud j sont définies par les relations de partage suivantes:

$$\psi_{j+1,2n}(t) = \sum_{k=-\infty}^{+\infty} h(k) \psi_{j,n}(t - 2^j k) \quad (2.33)$$

et

$$\psi_{j+1,2n+1}(t) = \sum_{k=-\infty}^{+\infty} g(k) \psi_{j,n}(t - 2^j k) \quad (2.34)$$

Les coefficients des paquets d'ondelettes se calculent par un algorithme de banc de filtres qui généralise la transformée en ondelettes discrète rapide.

Comme $\{\psi_{j,n}(t - 2^j k)\}_{k \in \mathbb{Z}}$ est orthonormée,

$$h(k) = \langle \psi_{j+1,2n}(t), \psi_{j,n}(t - 2^j k) \rangle, \quad g(k) = \langle \psi_{j+1,2n+1}(t), \psi_{j,n}(t - 2^j k) \rangle \quad (2.35)$$

$\{\psi_{j+1,2n}(t - 2^{j+1} k)\}_{k \in \mathbb{Z}}$ et $\{\psi_{j+1,2n+1}(t - 2^{j+1} k)\}_{k \in \mathbb{Z}}$ sont les bases orthonormées de deux espaces orthogonaux $W_{j+1,2n}$ et $W_{j+1,2n+1}$ tels que

$$W_{j+1,2n} \oplus W_{j+1,2n+1} = W_{j,n} \quad (2.36)$$

Ce partage récursif définit un arbre binaire de paquets d'ondelettes où chaque nœud père est subdivisé en deux espaces orthogonaux [Mal99]. La figure 2.6 montre les 8 paquets d'ondelettes $\psi_{j,n}$ à la profondeur $j = 3$, calculés à partir des filtres de Daubechies d'ordre 5.

2.4.3 Calcul des coefficients des paquets d'ondelettes

On notera C l'arbre de la projection du signal dans les sous espaces de W (C a la même structure d'arbre que W). On définit un paquet d'ondelettes $C_{j,n} = \{C_{j,n}(k)\}_{k \in \mathbb{Z}}$ comme étant la projection du signal dans le sous espace $W_{j,n}$ dont $\psi_{j,n}$ est une base orthonormée.

Pour tout nœud (j,n) de l'arbre des paquets d'ondelettes, on note les coefficients des paquets d'ondelettes d'une fonction $f(t)$ par:

$$C_{j,n}(k) = \langle f(t), \psi_{j,n}(t - 2^j k) \rangle \quad (2.37)$$

La décomposition en paquets d'ondelettes est obtenue à l'aide d'un algorithme pyramidal étendu (figure 2.3) par un calcul récursif :

$$C_{j+1,2n}(k) = \sum_l^{\wedge} h(2k-l) C_{j,n} \quad (2.38)$$

$$C_{j+1,2n+1}(k) = \sum_l^{\wedge} g(2k-l) C_{j,n} \quad (2.39)$$

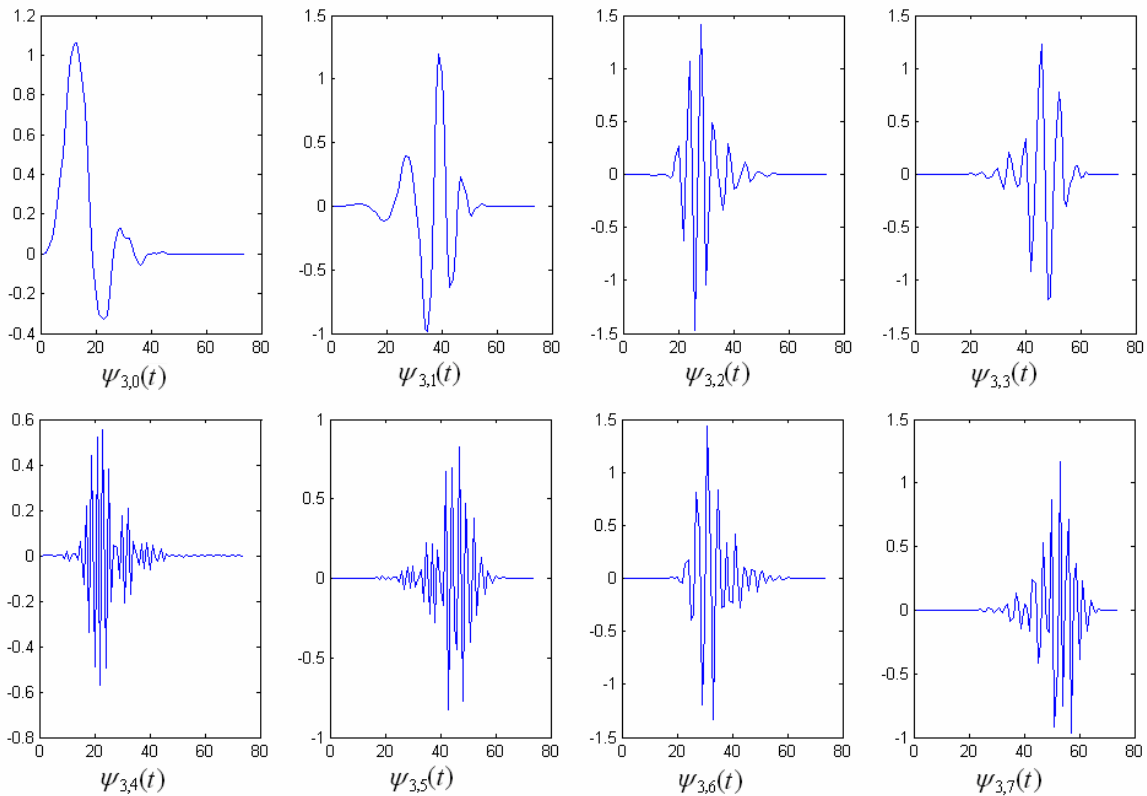


Figure 2.6. Paquets d'ondelettes calculés avec le filtre de Daubechies d'ordre 5, à la profondeur $j = 3$ de l'arbre des paquets d'ondelettes.

Ainsi le paquet $C_{j+1,2n}$ (respectivement $C_{j+1,2n+1}$) est obtenu par filtrage passe-bas (resp. passe haut) de $C_{j,n}$ par le filtre h (resp. g), suivi d'une décimation qui projette le signal sur le sous espace $W_{j+1,2n}$ (resp. $W_{j+1,2n+1}$).

La figure 2.7 reprend un signal synthétique à base de segments gaussiens du même type que celui de la figure 2.1, filtrés dans des bandes fréquentielles. La décomposition en paquets d'ondelettes est faite jusqu'au niveau 3. Les coefficients des paquets correspondant aux espaces $W_{2,0}$, $W_{3,2}$, $W_{3,3}$, $W_{2,2}$ et $W_{2,3}$ sont représentés (pour simplifier dans la suite $W_{j,n}$ désigne le paquet correspondant à l'espace $W_{j,n}$).

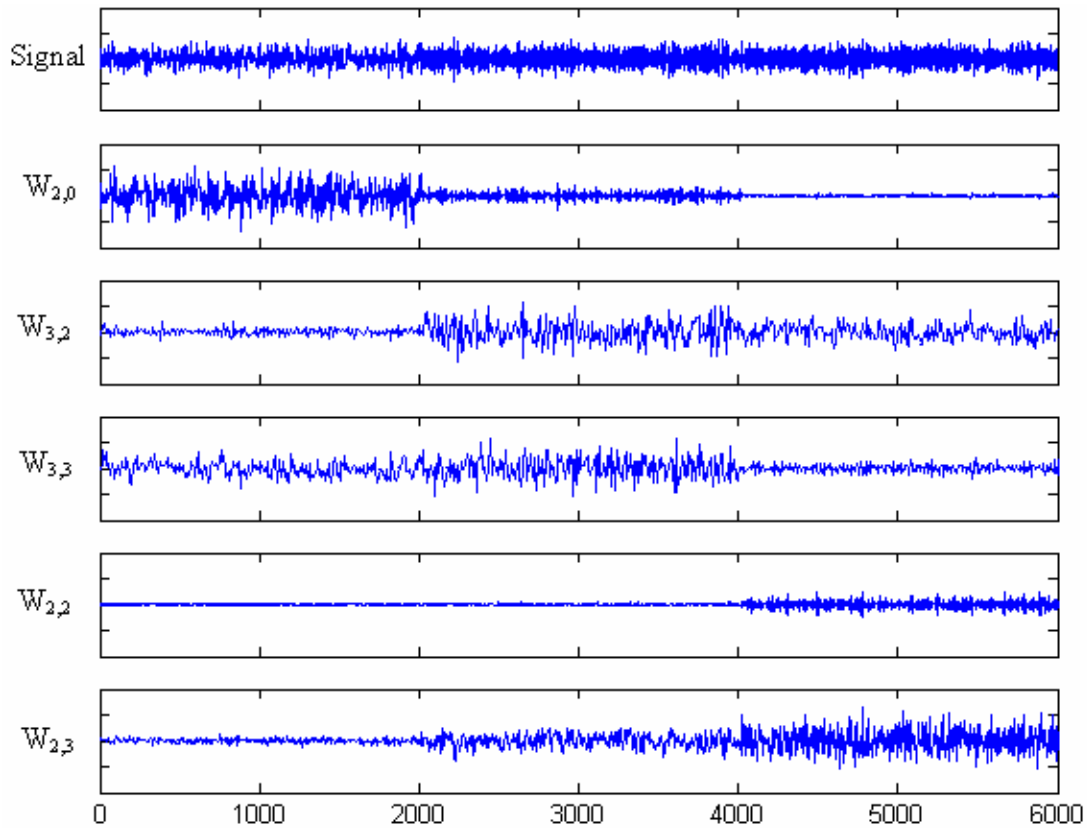


Figure 2.7. Décomposition d'un signal synthétique jusqu'au niveau 3. Les coefficients des paquets $W_{2,0}$, $W_{3,2}$, $W_{3,3}$, $W_{2,2}$ et $W_{2,3}$ sont représentés. Axes des abscisses : nombre des points. Axes des ordonnées : Unités arbitraires.

2.4.4 Arbre admissible

La décomposition totale d'un signal en paquets d'ondelettes fournit une représentation redondante. Chaque niveau de la décomposition contient toute l'information du signal. Ainsi seules certaines combinaisons de paquets constituent des représentations non redondantes du signal initial. On appelle arbre admissible tout arbre qui permet une reconstruction exacte du signal de départ. Il est représenté par un arbre binaire dont chaque nœud est à 0 ou 2 fils. L'arbre représenté sur la figure 2.8 est un arbre admissible.

Soient $\{j_i, n_i\}$ $1 \leq i \leq I$ les feuilles (les extrémités) d'un arbre binaire admissible. En appliquant le partage récursif (équation 2.36) le long des branches de cet arbre, on vérifie que les espaces $\{W_{j_i, n_i}\}_{1 \leq i \leq I}$ sont orthogonaux deux à deux et que leur somme vaut W_0 (figure 2.8):

$$W_0 = \bigoplus_{i=1}^I W_{j_i, n_i} \quad (2.40)$$

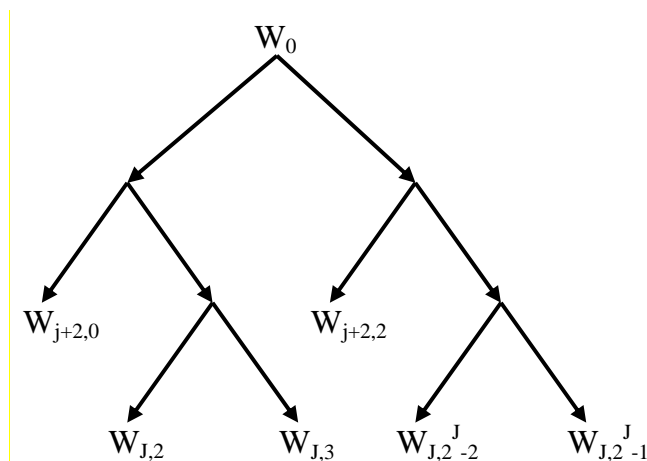


Figure 2.8. Exemple d'arbre binaire admissible de paquets d'ondelettes

La réunion des bases de paquets d'ondelettes correspondants $\{\psi_{j_i, n_i}(t - 2^{j_i} k)\}_{n \in \mathbb{Z}, l \leq i \leq L}$ définit, ainsi, une base orthogonale de W_0 .

Le nombre de bases orthogonales de paquets d'ondelettes de W_0 est égal au nombre des différents arbres admissibles qui est supérieur à $2^{2^{J-1}}$ bases orthogonales de paquets d'ondelettes contenues dans un arbre binaire de profondeur J [Mal99]. On retrouve évidemment, parmi la multitude des bases, celle de la décomposition discrète en ondelettes.

2.5 Sélection d'une meilleure base des paquets d'ondelettes

A chaque niveau de l'arbre complet, toute l'information du signal est représentée. Un découpage fréquentiel régulier correspond à chaque niveau de décomposition donné. Si on complète la décomposition jusqu'à avoir un seul point dans chaque paquet terminal, la fréquence est alors représentée par ce point. Toute l'information temporelle est alors cachée et on a une décomposition purement fréquentielle du signal.

Deux facteurs importants déterminent la sélection d'une meilleure base : le choix de la représentation sur laquelle s'appuiera la sélection d'une part, et le critère de sélection utilisé d'autre part. Le critère devra être adapté aux objectifs de traitement : dans un objectif de compression, il mesure le degré de concentration du signal sur une base. Dans un problème de classification supervisée, il sera fonction de la capacité à discriminer les classes. Dans un objectif de détection, il sera défini pour permettre une séparation optimale des différentes composantes fréquentielles du signal (cas des ruptures fréquentielles).

L'algorithme du choix de la meilleure base permet de représenter un signal sur une base d'ondelettes de manière optimale au sens d'une fonction de coût, comme par exemple l'entropie de Shannon.

Dans le cas des paquets d'ondelettes, la meilleure base est obtenue à partir de l'arbre de décomposition dyadique du signal. Le principe d'une telle décomposition est, comme on l'a déjà précisé, de diviser de manière itérative chaque paquet en deux nouveaux paquets obtenus par filtrages passe-bas et passe-haut sur la moitié de la bande fréquentielle. Ainsi à une itération i de la décomposition, chaque paquet couvre une fraction $\frac{1}{2^i}$ de la bande fréquentielle, d'où le terme de sous-bande qualifiant chaque paquet.

D'autres approches de sélection de base concernant les applications de débruitage, de classification supervisée et de détection de ruptures fréquentielles, seront également présentés dans la suite.

2.5.1 Les algorithmes fondamentaux de recherche de la meilleure base

La notion de meilleure base associée aux librairies de bases orthogonales a été introduite par Coifman Wicherhauser [Coif92]. Elle était alors utilisée dans le cadre de la compression des signaux. L'idée de cette méthode est de trouver la base sur laquelle l'information est la plus concentrée. Un critère additif est défini pour mesurer ce degré de concentration. L'optimum global est obtenu par une stratégie de recherche locale. Cette méthode présente un caractère de généralité dans la mesure où elle définit une stratégie de recherche de l'optimum applicable à tout critère correspondant à une mesure additive. Le coût fréquemment utilisé est l'entropie de Shannon. Pour une série discrète $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$, l'entropie de Shannon est définie par :

$$M(\mathbf{x}) = -\sum_{i=1}^N (x_i)^2 \log(x_i)^2 \quad (2.41)$$

avec $(x_i)^2 \log(x_i)^2 = 0$ si $x_i = 0$.

$M(\mathbf{x})$ reflète la concentration d'énergie dans la série discrète \mathbf{x} . Pour que l'algorithme soit rapide, $M(\mathbf{x})$ doit être additif comme c'est montré en (2.41) [Hit99].

De façon générale, soient, M un critère additif, \mathbf{x} un vecteur dans un sous-espace W , B une base orthogonale de W issue d'une librairie de bases, et $B.\mathbf{x}$ la séquence de \mathbf{x} projetée dans B . La meilleure base de $\mathbf{x} \in W$ relativement à M est alors définie comme étant la base pour laquelle $M(B.\mathbf{x})$ est optimal, (selon la signification du critère, M doit être maximisé ou minimisé).

2.5.2 Les critères de sélection

L'entropie est ici le reflet de la quantité d'information présente à chaque noeud de l'arbre contenant les coefficients de paquets d'ondelettes. Nous allons lister ci-dessous quatre formules entropiques différentes, \mathbf{x} étant le signal envisagé et (x_i) les coefficients de \mathbf{x} dans la base orthonormée des paquets d'ondelettes.

1- L'entropie de Shannon (non normalisée) est:

$$M(\mathbf{x}) = -\sum_i x_i^2 \log(x_i^2) \quad (2.42)$$

avec la convention que $0 \log(0) = 0$ lorsque $x_i = 0$.

2- La concentration avec la norme l^p avec $l \leq p$

$$M(\mathbf{x}) = \sum_i |x_i|^p = \|\mathbf{x}\|_p^p \quad (2.43)$$

3- Le critère d'énergie logarithmique [Rav98] :

$$M(\mathbf{x}) = \sum_i \log(|x_i|^2) \quad (2.44)$$

4- La limite d'entropie:

$$M(x_i) = \begin{cases} 1 & \text{si } |x_i| > \varepsilon \\ 0 & \text{si non} \end{cases} \quad (2.45)$$

D'autres définitions dérivées de l'entropie de Shannon peuvent exister selon l'application en cours.

Critères de similarité

Pour mesurer la similarité entre deux séquences \mathbf{x} et \mathbf{y} , plusieurs mesures peuvent exister.

a. Distance L_p

$$L_p(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^n (x_i - y_i)^p \right)^{\frac{1}{p}} \quad \text{avec } p \in [1, \infty[\quad (2.46)$$

Pour $p = 2$, elle correspond à la distance euclidienne qui est la plus utilisée.

b. Divergence de Kullback-Leibler

La divergence de Kullback-Leibler (K-L) entre deux distributions de probabilité f et g est aussi connue sous le nom d'entropie relative :

$$D(f, g) = \int_x f \log \frac{f}{g} dx \quad (2.47)$$

De plus, comme $D(f; g) \neq D(g; f)$, on définit la distance de Kullback-Leibler (K) (1951) [Kul51] comme :

$$K(f; g) = K(g; f) = D(f, g) + D(g, f) \quad (2.48)$$

C'est la meilleure mesure pour la recherche d'information dans les grandes bases de données.

c. Similarité par cosinus

Cette mesure de similarité généralement associée à un modèle vectoriel correspond au cosinus de l'angle α formé par les deux vecteurs \mathbf{x} et \mathbf{y} dans l'espace multidimensionnel.

$$\cos(\alpha) = \frac{\sum_{j=1}^n x_j \cdot y_j}{\sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2}} \quad (2.49)$$

2.5.3 Algorithme de Coifman Wicherhauser

A partir d'une décomposition complète en paquets d'ondelettes, la recherche de la meilleure base est effectuée du bas vers le haut, en comparant le coût de deux sous-bases possibles : la sous-base correspondant à un nœud père, et celle résultant de l'union de ses deux fils. Si le critère est minimum au niveau du père, on conserve celui-ci, sinon on remplace le père par ses deux fils.

En désignant par B une librairie de bases organisée en arbres binaires de profondeur maximale J , l'algorithme peut être résumé de la manière suivante:

- A_{J-1} est la meilleure base initiale constituée des sous-bases

$$\{B_{J,n}\}, \quad n = 0, \dots, 2^J - 1$$

- pour chaque niveau $j = J - 2$ à 0 , pour chaque nœud $n = 0, \dots, (2^j - 1)$, la meilleure sous-base $A_{j,n}$ est obtenue par:

$$A_{j,n} = \begin{cases} B_{j,n} & \text{si } M(B_{j,n} \mathbf{x}) \leq M(A_{j+1,2n} \mathbf{x}) + M(A_{j+1,2n+1} \mathbf{x}) \\ A_{j+1,2n} \oplus A_{j+1,2n+1} & \text{si non} \end{cases} \quad (2.50)$$

Le fait, d'une part, d'effectuer ce traitement à partir du bas de l'arbre en prenant comme base initiale la totalité de ses feuilles, et de retenir exclusivement comme sous-base les deux fils ou leur père d'autre part, garantit la convergence de l'algorithme vers l'obtention d'une base optimale admissible [Coif92, Hit98].

2.5.4 Sélection d'une meilleure base pour la compression [Coif92]

Dans le cadre de la compression des signaux, il est nécessaire de maximiser la concentration de l'énergie sur un petit nombre de vecteurs. Le critère sera une fonction de coût additive permettant de mesurer cette concentration de l'information, comme par exemple l'entropie de Shannon du signal. Il sera ensuite nécessaire d'adopter une stratégie de recherche de la base permettant d'optimiser ce critère. Elle est généralement basée sur l'algorithme développé dans le paragraphe précédent.

2.5.5 Sélection d'une meilleure base pour le débruitage

Bien que l'objectif final soit différent, la recherche de meilleures bases dans le cadre du débruitage utilise souvent les critères développés pour la compression. Les méthodes non linéaires d'estimation d'un signal à partir d'une mesure bruitée par un bruit gaussien (soft ou hard denoising) ont été développées par Donoho [Don94], [Don95], initialement à partir de la transformée en ondelettes discrète du signal. Elles utilisent le pouvoir de compression de cette transformation pour séparer le bruit (réparti sur l'ensemble des coefficients d'ondelettes avec de faibles amplitudes) et le signal utile (caractérisé la plupart du temps par quelques coefficients de forte amplitude) en réalisant un seuillage adapté au niveau de bruit. Les performances de la méthode sont d'autant meilleures que l'information utile du signal est concentrée sur un nombre réduit de coefficients. Il est donc naturel de rechercher des bases maximisant la concentration du signal, ce qui peut être réalisé avec tout critère de type entropique. Toutefois, Donoho a défini un critère plus spécifiquement adapté au débruitage : la mesure utilisée n'est plus l'entropie définie précédemment, mais une fonction de la dimension théorique du signal calculée sur la base d'un seuil optimal pour le débruitage [Don94] [Don95a].

2.5.6 Sélection d'une meilleure base pour la classification supervisée

La méthode de Saito et Coifman [Sai94] est inspirée de la méthode développée pour la compression. Le but est de trouver la base qui maximise la séparabilité entre des classes d'apprentissage. Le critère utilisé sera donc un critère de discrimination évaluant le pouvoir discriminant de chacun des nœuds représentant une sous-base potentielle, vis-à-vis des classes à séparer. L'entropie relative ou « cross entropy » constitue une mesure de distance entre les distributions temporelles d'énergie de deux séquences \mathbf{x} et \mathbf{y} . Elle est définie par [Sai94]:

$$D(\mathbf{x}, \mathbf{y}) = \sum_i p_i \log\left(\frac{p_i}{q_i}\right) \text{ où } p_i = \frac{|x_i|^2}{\|\mathbf{x}\|^2} \text{ et } q_i = \frac{|y_i|^2}{\|\mathbf{y}\|^2} \quad (2.51)$$

Cette entropie constitue le critère à optimiser. Elle correspond à la divergence de Kullback entre les distributions de deux séquences \mathbf{x} et \mathbf{y} représentatives de deux classes.

Les étapes de l'algorithme sont les suivantes :

- Les classes d'apprentissage sont chacune représentée par un arbre moyen. Celui-ci contient pour chaque noeud, une séquence moyenne des carrés des coefficients des individus de la classe.
- Si on a deux classes, l'arbre de décision est alors directement calculé en utilisant l'information de Kullback. Sinon, le critère est obtenu en sommant les mesures calculées sur chacun des couples de classes.
- Le critère étant additif, on applique la stratégie explicitée pour la compression, on trouve alors la base qui sera la plus discriminante pour ces classes de signaux.

Afin de réduire davantage la dimension du problème, une étape supplémentaire consiste à ne retenir qu'un nombre restreint de fonctions de base correspondant aux coefficients les plus discriminants.

2.5.7 Sélection de la meilleure base pour la détection des ruptures fréquentielles

Cette approche correspond tout à fait à la problématique de segmentation et d'isolation d'événements spécifiques dans un enregistrement, lorsque ces événements se caractérisent par leur contenu fréquentiel. Dans [Hit99], le but est de détecter des ruptures fréquentielles dans des signaux multicomposantes stationnaires par morceaux. Les auteurs proposent un algorithme qui consiste à chercher parmi toutes les bases possibles obtenues après décomposition en paquets d'ondelettes, celle qui représente le banc de filtres le plus adapté au signal.

La méthode présentée ici est différente de celle présentée pour la compression. Le critère entropique réalise une représentation concentrée du signal. La base choisie est alors reliée à une localisation fréquentielle très précise au détriment de la localisation temporelle. Le choix du critère énergétique est donc approprié à ces objectifs.

La sélection de la meilleure base s'effectue en trois étapes :

- a. Localiser les nœuds de la base correspondant à une activité significative du signal : un critère énergétique décide si le nœud est d'énergie significative en comparant l'énergie à un seuil s . On obtient alors un nouvel arbre A de même structure:

$$A_{j,n} = \begin{cases} 1 & \text{si } \text{energie}(C_{j,n}) > s \\ 0 & \text{si } \text{non} \end{cases} \quad (2.52)$$

- b. Indiquer le nombre des composantes présentes dans chaque paquet.
- c. Sélectionner les paquets monocomposantes de plus grande résolution temporelle, c'est-à-dire les plus proches de la racine de l'arbre.

La figure 2.9 montre bien les trois étapes de sélection de la meilleure base. La première étape consiste à calculer les énergies des paquets, mettre à 1 les paquets dont les coefficients ont une énergie supérieure à un seuil s , et mettre à 0 les autres (figure 2.9-a). La deuxième étape consiste à identifier les paquets multicomposantes : ce sont les paquets dont le nombre de composantes est strictement supérieur à 1. Le nombre de composantes d'un paquet est égal à la somme du nombre de composantes de ses fils (figure 2.9-b). Il suffit alors de parcourir l'arbre des feuilles vers la racine pour obtenir ce nombre de composantes.

La dernière étape est de sélectionner les paquets non multi-composantes de plus haut niveau, c'est-à-dire ayant une meilleure résolution temporelle. Il s'agit des nœuds à 0 ou à 1 ayant un père supérieur ou égal à 2, ce qui est bien représenté sur la figure 2.9-c.

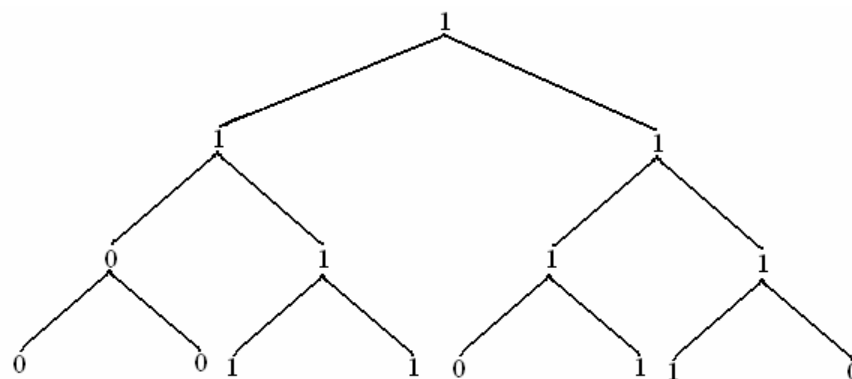
Les performances de cet algorithme vont dépendre principalement de [Hit99]:

- la décomposition utilisée et donc de la capacité des filtres à concentrer l'énergie d'une composante au sein d'un seul paquet.
- la pertinence du choix du seuil énergétique.

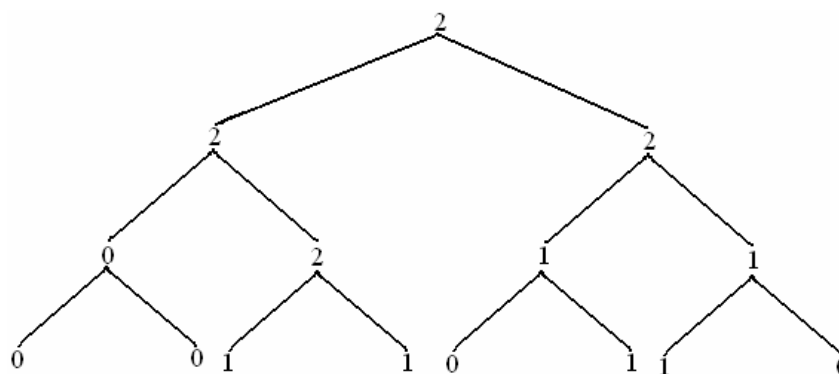
Dans [Hit99], Hitti a proposé deux approches de calcul du seuil pour le choix des paquets significatifs. La connaissance de l'énergie *a priori* des paquets d'ondelettes aide bien à calculer le seuil. La somme des énergies des paquets de chaque niveau j $EnsSP_j$ est estimée et l'énergie des paquets est comparée à un seuil dépendant du niveau j :

$$S_j = \frac{EnsSP_j}{2^j} \quad (2.53)$$

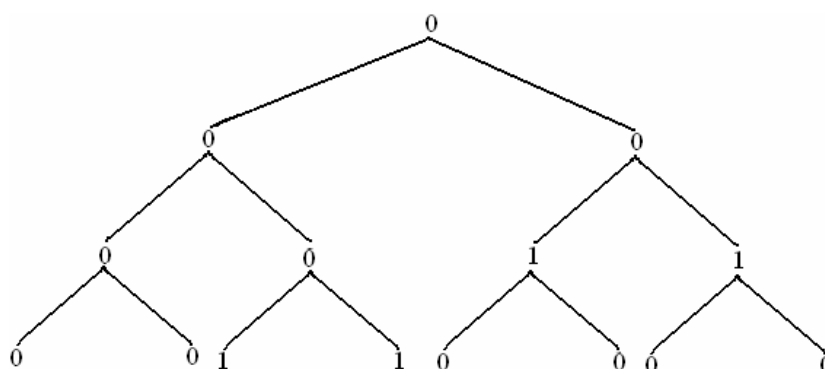
Les filtres de décomposition sont utilisés de telle sorte que l'énergie moyenne des paquets de chaque niveau soit égale à celle du signal original. Pour cela, les énergies significatives des événements sont concentrées dans certains paquets à un certain niveau, d'où l'interprétation du calcul du seuil $EnsSP_j$. La figure 2.10 montre les paquets de la meilleure base choisie par cette méthode. Elle montre bien la distinction entre les différents événements : le signal original est formé de trois segments de même amplitude à fréquences différentes.



a- arbre obtenu après seuillage de l'énergie des paquets.



b- nombre de composantes par paquets



c- La meilleure base est constituée des nœuds à 1.

Figure 2.9. Etapes de la sélection d'une meilleure base pour la détection des ruptures fréquentielles en se basant sur l'algorithme de Hitti.

2.6 Conclusion

La détection et la classification doivent être appliquées sur des représentations spécifiques du signal original. Deux types de représentations ont été présentés dans ce chapitre : la première représentation est fondée sur la décomposition discrète en ondelettes. La deuxième représentation est fondée sur la décomposition en paquets d'ondelettes.

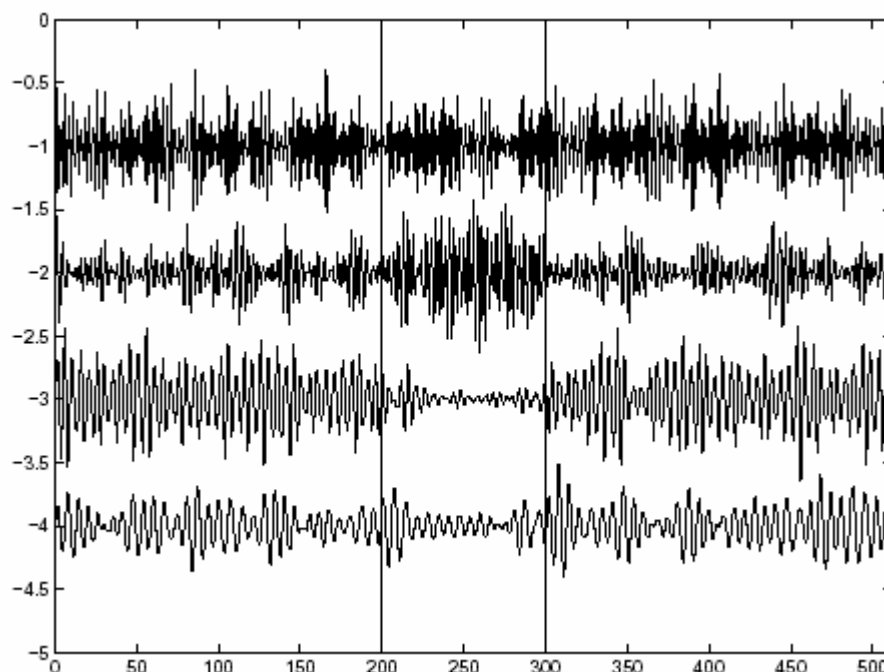


Figure 2.10. Les signaux reconstruits à partir des 4 paquets sélectionnés (voir figure 4.9) du signal original. Les lignes verticales indiquent les positions des vrais instants de changement d'après [Hit98].

A partir du signal original, nous pouvons obtenir plusieurs séquences représentant des coefficients de détail et des coefficients correspondant à des paquets d'ondelettes. La méthode de décomposition multiéchelles a été décrite dans ce chapitre ainsi que la méthode de décomposition en paquets d'ondelettes. Les différentes stratégies de choix de la meilleure base ont été présentées. Les objectifs du choix de la meilleure base étant divers, différents critères ont été décrits.

Une fois le signal modélisé par décomposition en paquets d'ondelettes, nous pouvons appliquer une méthode de choix d'une meilleure base pour des objectifs de détection des événements. Le chapitre suivant est consacré à l'analyse du critère de distance de Kullback Leibler pour montrer son efficacité à mettre en évidence la présence de ruptures dans un paquet donné, et par suite proposer un algorithme de sélection d'une meilleure base fondé sur ce critère.

Chapitre 3

Sélection d'une meilleure base pour la détection des transitoires

3.1 Introduction

Notre proposition de nouvelle approche de sélection de la meilleure base consiste à identifier les paquets révélant au mieux l'existence d'événements dans un signal donné. Nous avons vu que dans la décomposition en paquets d'ondelettes, chaque paquet représente l'information relative à tout le signal dans une bande de fréquence donnée. La sous base relative à un paquet peut donc être associée à un filtre et la base à un ensemble de filtres réalisant une partition totale ou non de l'axe fréquentiel. L'idée à la base de l'algorithme proposé consiste à rechercher, parmi toutes les bases possibles, celle qui représente le banc de filtres le mieux adapté au signal en terme de mise en évidence des événements existant dans le signal de départ.

La volonté d'obtenir des paquets d'ondelettes qui nous donnent des informations concernant la présence d'événements nous permet ensuite d'appliquer un algorithme de détection directement sur les paquets sélectionnés. De plus, la précision du détecteur en termes d'instant de changement dépendra de la résolution temporelle des coefficients des paquets d'ondelettes considérés. Si l'on considère l'arbre complet issu d'une décomposition, les nœuds sont caractérisés par une résolution fréquentielle croissante et par une résolution temporelle décroissante lorsqu'on descend dans cet arbre.

Ayant défini les objectifs de détermination de la meilleure base capable de mettre en évidence des événements spécifiques dans un enregistrement, se pose alors la question du critère à

utiliser et de la stratégie de sélection de la meilleure base qui lui est associée. La méthode que nous proposons utilise comme critère la distance de Kullback Leibler (K-L), compte tenu de sa capacité à mettre en évidence les différences de caractéristiques statistiques entre catégories d'événements.

3.2 Critère de sélection

Dans la description des critères de sélection de meilleures bases du chapitre précédent, le coût défini par la fonction additive M est lié à la distribution de l'énergie du signal sur les différents nœuds d'un arbre de décomposition. L'optimisation de cette fonction amène à une représentation du signal minimisant la perte d'information. Du fait de la définition même de cette fonction de coût, le choix de la meilleure base est bien adapté à des problèmes liés à la compression. Il n'en est évidemment de même pour un problème de détection. En effet, dans ce cadre, nous avons besoin d'une fonction qui choisit parmi tous les paquets d'ondelettes, ceux qui ont la capacité à mettre en évidence les changements d'événements dans les enregistrements. Or la distance de K-L a déjà été utilisée dans le passé comme mesure discriminante; en particulier pour la comparaison d'images [Kul51], [Do02], ou encore pour des problèmes de classification [Coi92].

Pour une classe donnée de signaux, en particulier pour les signaux EMG utérin, certains paquets représentent bien l'existence d'événements spécifiques comme les contractions, les ondes d'Alvarez, les mouvements actifs fœtaux etc., alors que d'autres ne représentent que du bruit. La figure 3.1 montre la décomposition d'un signal EMG utérin appartenant à la base REELSIM. Il est formé de 10 événements : 3 contractions, 3 mouvements actifs fœtaux, 2 ondes d'Alvarez et 2 ondes LDBF. Ces événements ont été artificiellement reliés et bruités, avec un rapport signal sur bruit de 10 dB. Seuls les coefficients des paquets 3, 8, 6 et 14 sont représentés. Il est clair que les paquets 3 et 8 sont capables de différencier les événements de la ligne de base, alors que les paquets 6 et 14 ne donnent aucune indication sur la présence d'événements. Cet exemple illustre l'idée de choisir un ensemble de paquets pertinents pour détecter les événements existants et d'éliminer les autres. Ce qui se pose alors comme question est le choix d'un critère pour sélectionner les paquets qui traitent du problème de détection des événements, sans prendre en compte pour l'instant l'exactitude des instants des changements. La figure 3.1 montre qu'il y a une différence importante entre les distributions statistiques des coefficients sur chacun des nœuds. Dans les paquets où il n'y a pas de

distinction claire sur la présence d'événements, les distributions statistiques sont sensiblement les mêmes. On montrera que les distributions ont un caractère gaussien pour des signaux réels. Par contre on montrera que ce caractère gaussien n'est plus vérifié pour les nœuds faisant bien ressortir les événements, d'où l'idée de s'orienter vers un critère qui mette en évidence ces différences de distributions statistiques.

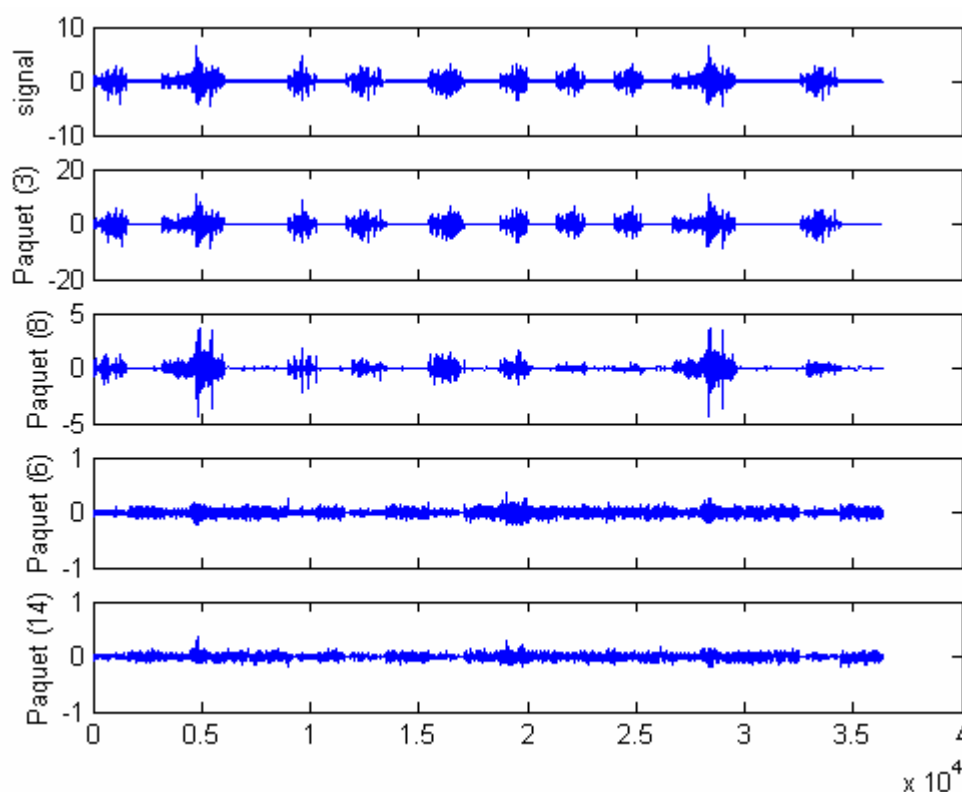


Figure 3.1. Décomposition en paquets d'ondelettes d'un signal EMG utérin formé de 10 événements. Les coefficients des paquets numéro 3, 8, 6 et 14 sont représentés. Axes des abscisses : nombre des points. Axes des ordonnées : Unités arbitraires.

De façon plus précise, quelle que soit la fenêtre d'observation définie dans un nœud n'exprimant que du bruit, la distribution des amplitudes (ou des énergies instantanées) de la série temporelle contenue dans la fenêtre sera du même type (gaussienne). Par contre, pour des nœuds exprimant bien la présence d'événements, la distribution liée à la série temporelle contenue dans la fenêtre dépendra de la position de cette fenêtre (selon que celle-ci contient un événement ou ne contient que du bruit). Un indicateur de distance est la distance de

Kullback-Leibler [Cov91] $D(f, g) = \int_x f \log \frac{f}{g} dx$ ou, lorsque les deux fonctions ne diffèrent

que par leur paramètre θ :

$$D(f(X, \theta_i), g(X, \theta_j)) = \int f(x, \theta_i) \log \frac{f(x, \theta_i)}{g(x, \theta_j)} dx. \quad (3.1)$$

Cette expression calcule bien la distance entre deux distributions de densités de probabilité $f(X, \theta_i)$ et $f(X, \theta_j)$, et met en évidence l'augmentation de distance quand les deux distributions sont différentes, alors que cette même distance est théoriquement nulle lorsque les distributions sont les mêmes. Cette propriété de la distance de Kullback Leibler va nous permettre de l'utiliser comme indicateur de la présence de ruptures dans les différents nœuds de l'arbre de décomposition.

A partir de la formulation générale de la distance de Kullback Leibler donnée par l'équation 3.1, il est possible d'envisager une estimation de cette distance à partir de séquences temporelles limitées pour lesquelles la distribution des éléments de la séquence est accessible à l'estimation.

Dans le cadre de l'étude des signaux électromyographiques utérins, l'estimation de cette distance, puis son utilisation dans le but de sélectionner les meilleurs paquets d'ondelettes, passe par un certain nombre d'étapes :

- Etape 1 : démontrer que la distribution des éléments d'une séquence donnée varie en fonction du contenu de la séquence. Cela revient à montrer que la distribution des coefficients d'un paquet d'ondelettes donné change selon que ce paquet représente du bruit ou un événement spécifique utile à l'étude (contractions, ondes d'Alvarez etc.). Cette étape est développée dans le paragraphe 3.3, dans lequel on montrera que les coefficients suivent une loi gaussienne généralisée pour les événements de l'EMG utérin, alors qu'ils se résument à une loi normale pour le bruit (absence d'événements dans l'enregistrement).
- Etape 2 : estimer la distance de Kullback Leibler. Une fois les densités de probabilité connues et leurs paramètres estimés, il est possible de remplacer leur expression dans l'équation 3.1, et ainsi d'obtenir une estimation de la distance de Kullback Leibler entre deux séquences comprenant ou non des événements utiles. Cette estimation est présentée dans le paragraphe 3.4.
- Etape 3 : estimer la distribution de la distance de Kullback Leibler. Ayant un outil d'estimation de la distance de Kullback Leibler entre deux séquences, l'idée est

d'utiliser à présent cet outil afin de sélectionner, parmi l'ensemble des paquets d'ondelettes décomposant un enregistrement d'EMG utérin, ceux qui mettent le mieux en évidence les événements utiles.

L'approche que nous avons utilisée prend comme base la séquence des coefficients obtenus au niveau d'un nœud de l'arbre de décomposition en paquets d'ondelettes.

L'idée est de découper cette séquence en fenêtres successives de largeur suffisante pour estimer correctement les paramètres de la distribution des coefficients, mais de largeur limitée afin de pouvoir contenir soit uniquement du bruit, soit uniquement une partie d'événement utile.

Dans le cas d'un nœud ne représentant que du bruit, toutes les fenêtres représenteront le même type de séquence (bruit). Ainsi, si l'on calcule les distances de Kullback Leibler entre tous les couples de fenêtres de ce nœud, on obtiendra un ensemble de distances dont la valeur théorique est nulle.

La distribution des distances obtenues sera alors représentative de la distribution de l'estimateur de la distance de Kullback Leibler en l'absence d'événements.

Par contre, dans le cas d'un nœud représentatif de la présence d'événements, la distribution des distances obtenues devrait s'éloigner significativement de la distribution en l'absence d'événements.

Le paragraphe 3.5 est dédié à l'étude de la distribution de la distance de Kullback Leibler en l'absence d'événements. Il propose une première approche de la modélisation de cette distribution à partir d'une distribution exponentielle. Il montre ensuite qu'en présence d'événements, la distribution obtenue n'a plus rien à voir avec ce modèle exponentiel.

- Etape 4 : sélectionner la meilleure base de paquets d'ondelettes. L'étape précédente fournit un outil de discrimination entre les nœuds de l'arbre de décomposition contenant des événements et ceux ne contenant que du bruit. L'étape suivante consiste à utiliser une base d'apprentissage pour estimer les distributions de la distance de Kullback Leibler à chacun des nœuds de l'arbre, et à comparer la distribution obtenue à la distribution théorique en l'absence d'événements (étape 3).

Les paquets retenus seront ceux pour lesquels la distribution obtenue s'éloignera le plus de la distribution précédente.

Afin de pouvoir évaluer cette distance entre distributions aussi bien en présence qu'en l'absence de modèle, nous avons cette fois choisi d'utiliser le D_{\max} de Kolmogorov Smirnov (paragraphe 3.6.1 et 3.6.2).

- Etape 5 : réduire la redondance. L'arbre de décomposition est hautement redondant. La sélection de paquets pertinents à l'étape précédente peut mener à des paquets contenant la même information. La dernière étape consiste alors à éliminer les paquets redondants pour ne conserver qu'un minimum d'entre eux contenant des informations complémentaires. Cette dernière phase du processus de sélection de paquets est décrite au paragraphe 3.6.3.

3.3 Densité gaussienne généralisée pour les coefficients des paquets d'ondelettes

Des études antérieures [Mou99], [Sha95], [Wou99] montrent qu'une bonne approximation de la distribution des coefficients des paquets d'ondelettes dans une sous bande particulière peut être obtenue par une variation adaptative des deux paramètres d'une densité gaussienne généralisée (GGD), qui est définie comme suit [Do02]:

$$f(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta} \quad (3.2)$$

où $\Gamma(\cdot)$ est la fonction gamma : $\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt$, $z > 0$. α (paramètre d'échelle) est relié

à la largeur de la distribution (écart type) alors que β est un paramètre de forme. La GGD est une distribution gaussienne lorsque $\beta = 2$.

3.3.1 Estimation des paramètres d'échelle et de forme α et β

Plusieurs approches ont été adoptées pour estimer α et β [Do02], [Var89], [Sha95]. Récemment, *Do* et *Vetterli* ont proposé une estimation basée sur le principe du maximum de vraisemblance [Do02]. Soit $L(\mathbf{x}; \alpha, \beta)$ la fonction de log-vraisemblance d'une observation $\mathbf{x} = (x_1, \dots, x_N)$ où les différents échantillons x_i sont considérés indépendants, on a :

$$L(\mathbf{x}; \alpha, \beta) = \log \prod_{i=1}^N f(x_i; \alpha, \beta) \quad (3.3)$$

où l'expression de f est donné par la GGD. Les paramètres α et β devront être estimés. Le principe du maximum de vraisemblance nous conduit aux équations ci-dessous :

$$\left\{ \begin{array}{l} \frac{\partial L(x; \alpha, \beta)}{\partial \alpha} = -\frac{N}{\alpha} + \sum_{i=1}^N \frac{\beta |x_i|^\beta \alpha^{-\beta}}{\alpha} = 0 \\ \frac{\partial L(x; \alpha, \beta)}{\partial \beta} = \frac{N}{\beta} + \frac{N\Pi(1/\beta)}{\beta^2} - \sum_{i=1}^N \left(\frac{|x_i|}{\alpha} \right)^\beta \log \left(\frac{|x_i|}{\alpha} \right) = 0 \end{array} \right. \quad (3.4)$$

où $\Pi(\cdot)$ est la fonction digamma [Abr70], i.e., $\Pi(z) = \Gamma'(z)/\Gamma(z)$.

Ce système d'équations admet une solution unique appelée l'estimée au sens du maximum de vraisemblance du couple de paramètres (α, β) . Quand on fixe le paramètre $\beta > 0$, l'estimée de α est alors donnée par :

$$\hat{\alpha} = \frac{\beta}{N} \left(\sum_{i=1}^N |x_i|^\beta \right)^{1/\beta} \quad (3.5)$$

L'estimée du paramètre de forme $\hat{\beta}$ est la solution de l'équation transcendante ci-dessous :

$$1 + \frac{\Pi(1/\hat{\beta})}{\hat{\beta}} - \frac{\sum_{i=1}^N |x_i|^{\hat{\beta}} \log |x_i|}{\sum_{i=1}^N |x_i|^{\hat{\beta}}} + \frac{\log \left(\frac{\hat{\beta}}{N} \sum_{i=1}^N |x_i|^{\hat{\beta}} \right)}{\hat{\beta}} = 0 \quad (3.6)$$

Cette équation est fortement et clairement non-linéaire. Cependant, elle peut être résolue grâce à l'algorithme de Newton-Raphson [Kay93], [Do02] qui est détaillé à l'annexe B.

Le paramètre d'échelle α est alors obtenu en réinjectant la valeur estimée de β dans l'équation donnant α (équation 3.5). Ainsi, en suivant la méthode développée à l'annexe B on arrive à estimer à la fois les paramètres d'échelle et de forme.

Dans ce qui suit, nous allons vérifier la validité de l'utilisation de ce type de distribution des coefficients des paquets d'ondelettes dans le cas de l'EMG utérin.

3.3.2 Distribution des coefficients des paquets d'ondelettes dans le cas de l'EMG utérin

Le but de cette section est de vérifier que les événements (Contractions, Mouvements Actifs Fœtaux, ondes d'Alvarez, ondes LDBF) et les enregistrements réels des signaux femmes suivent une loi gaussienne généralisée GGD, ainsi que leurs coefficients dans les paquets d'ondelettes de l'arbre de décomposition. L'expression de la densité gaussienne généralisée est donnée par l'équation 3.2 rappelée ci-dessous:

$$f(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta}$$

L'idée est d'estimer les valeurs de α et β en utilisant la méthode du maximum de vraisemblance expliquée dans le paragraphe précédent. Si $\beta = 2$ et $\alpha = \sigma\sqrt{2}$ (σ est l'écart type) alors le signal (ou l'événement de l'EMG utérin) suit une loi normale.

Dans une première étape, les valeurs de α et β ont été estimées pour l'ensemble d'apprentissage G1REEL1 (Annexe A) composé de 50 événements de contractions, ondes Alvarez, MAF, ondes LDBF et de 50 segments de bruit réel.

Les valeurs de α et β ont été estimées pour chaque type d'événement de deux manières. La première est de concaténer l'ensemble des éléments de chaque type d'événement puis d'estimer α et β une seule fois. La deuxième est d'estimer α et β pour chaque élément puis de calculer leur moyenne et leur écart type.

Il est ensuite nécessaire d'effectuer la comparaison entre les distributions des événements et la loi Gaussienne Généralisée GGD. Cette comparaison est faite en utilisant le test de Kolmogorov Smirnov (K-S) [Sap90]. La statistique de test de K-S D_{max} , résultant du calcul de la différence maximale entre les deux fonctions de répartition de la loi expérimentale des événements et de la loi théorique (GGD), est calculée pour décider si un événement suit une loi GGD ou pas.

Le tableau 3.1 résume les valeurs estimées de α , β et les valeurs calculées de D_{max} lors de la concaténation, alors que le tableau 3.2 résume les moyennes et les écart types de ces valeurs lors de leur estimation pour chaque élément. Les tableaux 3.1 et 3.2 contiennent en plus la probabilité $P_{D_{max}} = P(D > D_{max})$, probabilité que la statistique de Kolmogorov Smirnov

soit supérieure à D_{max} [Sap90]. La valeur de $P_{D_{max}}$ est naturellement influencée par l'échantillon utilisé.

	α	β	D_{max}	$P_{D_{max}}$
CT	1.0383	1.3303	0.0038	0.2
Alvarez	1.1947	1.5473	0.0081	0.1360
MAF	0.9557	1.2337	0.0119	0.0921
LDBF	0.9868	1.2711	0.0033	0.2
Bruit	1.3591	1.7529	0.0195	0.03

Tableau 3.1. Valeurs de α , β et D_{max} pour les événements de l'EMG utérin avec les segments du bruit réel (événements concaténés), $P_{D_{max}} = P(D > D_{max})$.

	$\bar{\alpha}$	σ_{α}	$\bar{\beta}$	σ_{β}	\bar{D}_{max}	$\sigma_{D_{max}}$	$\bar{P}_{D_{max}}$
CT	1.0425	0.2246	1.3850	0.3048	0.0385	0.0166	0.0895
Alv	1.1855	0.3367	1.7676	0.7583	0.0652	0.0254	0.1122
MAF	0.9742	0.4513	1.4679	0.7250	0.1019	0.0396	0.0887
LDBF	0.9910	0.2260	1.3246	0.3163	0.0312	0.0135	0.0617
Bruit	1.3094	0.2346	1.8582	0.5401	0.0510	0.0219	0.0888

Tableau 3.2. Valeurs de α , β et D_{max} pour les événements de l'EMG utérin avec les segments du bruit réel (événement par événement), $P_{D_{max}} = P(D > D_{max})$.

Dans le cas des événements concaténés, c'est la longueur totale des événements qui est considérée (plus de 10000 pour chaque type d'événement), alors que dans l'autre cas c'est la longueur moyenne des événements de chaque type qui est prise en considération (d'où un nombre de points beaucoup plus faible). Les écarts types, les valeurs de D_{max} et les probabilités $P_{D_{max}}$ montrent que les événements suivent bien une loi gaussienne généralisée.

Dans le cas du bruit réel, les segments suivent la GGD avec $\alpha \approx 1.3591$ et $\beta \approx 1.7529$, ce qui est très proche de la loi normale.

La figure 3.2 présente les histogrammes des événements concaténés de chaque type (définis à l'ensemble d'apprentissage G1REEL1 de la base REEL1) avec les tracés correspondants des fonctions des densités de GGD.

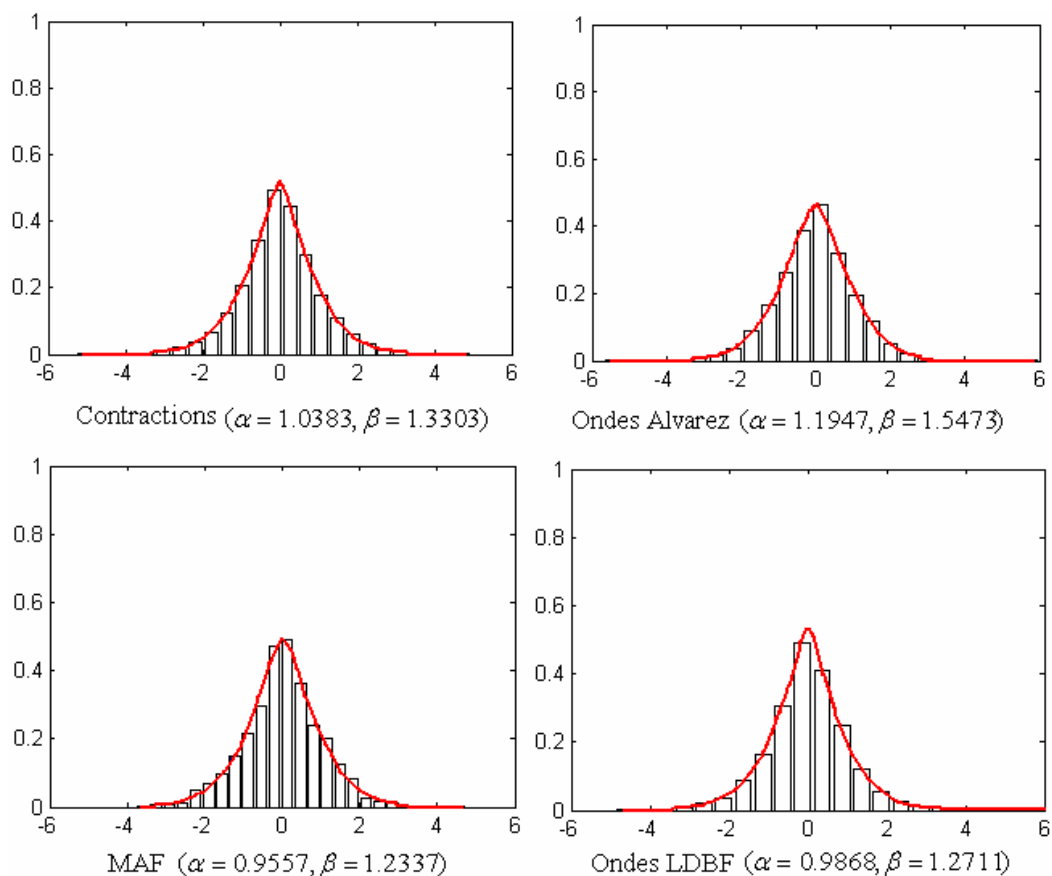


Figure 3.2. Histogrammes des événements de l'EMG utérin tracés avec les GGD associées, pour les événements concaténés de chaque type.

La figure 3.3 présente l'histogramme des segments concaténés du bruit réel définis dans l'ensemble d'apprentissage G1REEL1 de la base REEL1 et la distribution de la loi normale. On voit bien que cet histogramme est proche de la loi normale.

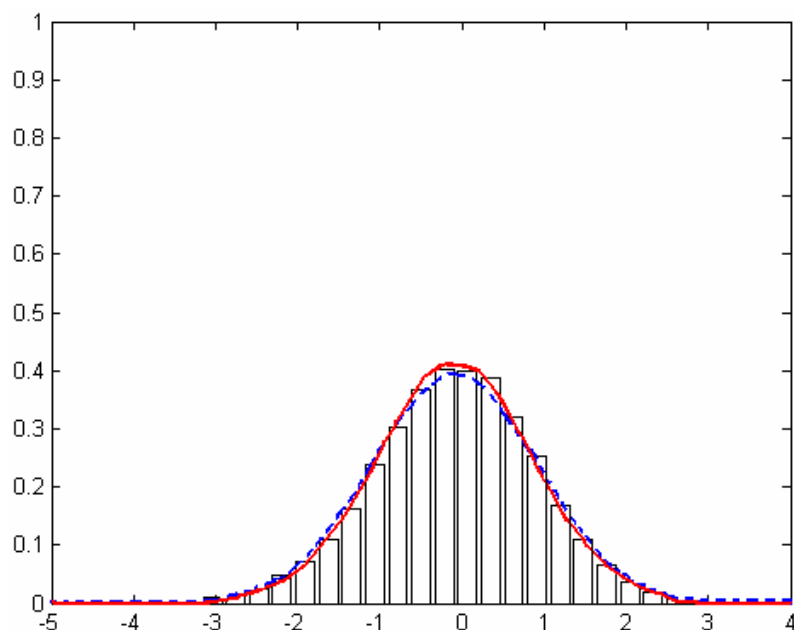


Figure 3.3. Histogramme du bruit réel avec la loi GGD associée aux valeurs $\alpha = 1.3591$ et $\beta = 1.7529$ (trait plein) et la loi normale (trait pointillé).

En complément de ce qui précède, les paramètres α et β ainsi que la statistique de test de K-S D_{max} sont calculés pour chaque signal femme défini à la base REEL2. Cette base est formée de 20 enregistrements réels de longue durée. Le tableau 3.3 résume les moyennes et les écarts types des paramètres estimés ainsi que des valeurs de D_{max} calculées avec $\bar{P}_{D_{max}} = P(D > D_{max}) = 0.1675$. Là encore on peut déduire que les amplitudes des signaux femmes suivent une distribution gaussienne généralisée GGD.

$\bar{\alpha} = 0.2384$	$\sigma_{\alpha} = 0.1427$
$\bar{\beta} = 0.6232$	$\sigma_{\beta} = 0.1265$
$\bar{D}_{max} = 0.00323$	$\sigma_{D_{max}} = 0.0055$
$\bar{P}_{D_{max}} = 0.1675$	

Tableau 3.3. Moyennes et écarts types de α , β et D_{max} de 20 signaux femmes de longue durée.

Pour ce qui concerne maintenant la loi suivie par les coefficients issus d'une décomposition en paquets d'ondelettes, cette décomposition est une transformation linéaire puisqu'elle est

réalisée par des filtrages successifs du signal. Elle conduit alors à des coefficients qui possèdent les mêmes propriétés statistiques que le signal de départ [Rav98]. Alors les coefficients des paquets d'ondelettes des événements utérins et des signaux femmes suivent une loi gaussienne généralisée. Le signal présenté à la figure 3.4.a en est une illustration. Il est décomposé en paquets d'ondelettes, les coefficients des paquets 6 et 7 et leurs histogrammes associés sont présentés.

De ce qui précède on peut conclure que chaque événement contenu dans un enregistrement d'EMG utérin peut être statistiquement modélisé par une densité gaussienne générale GGD, alors que le bruit associé peut être décrit par une distribution gaussienne.

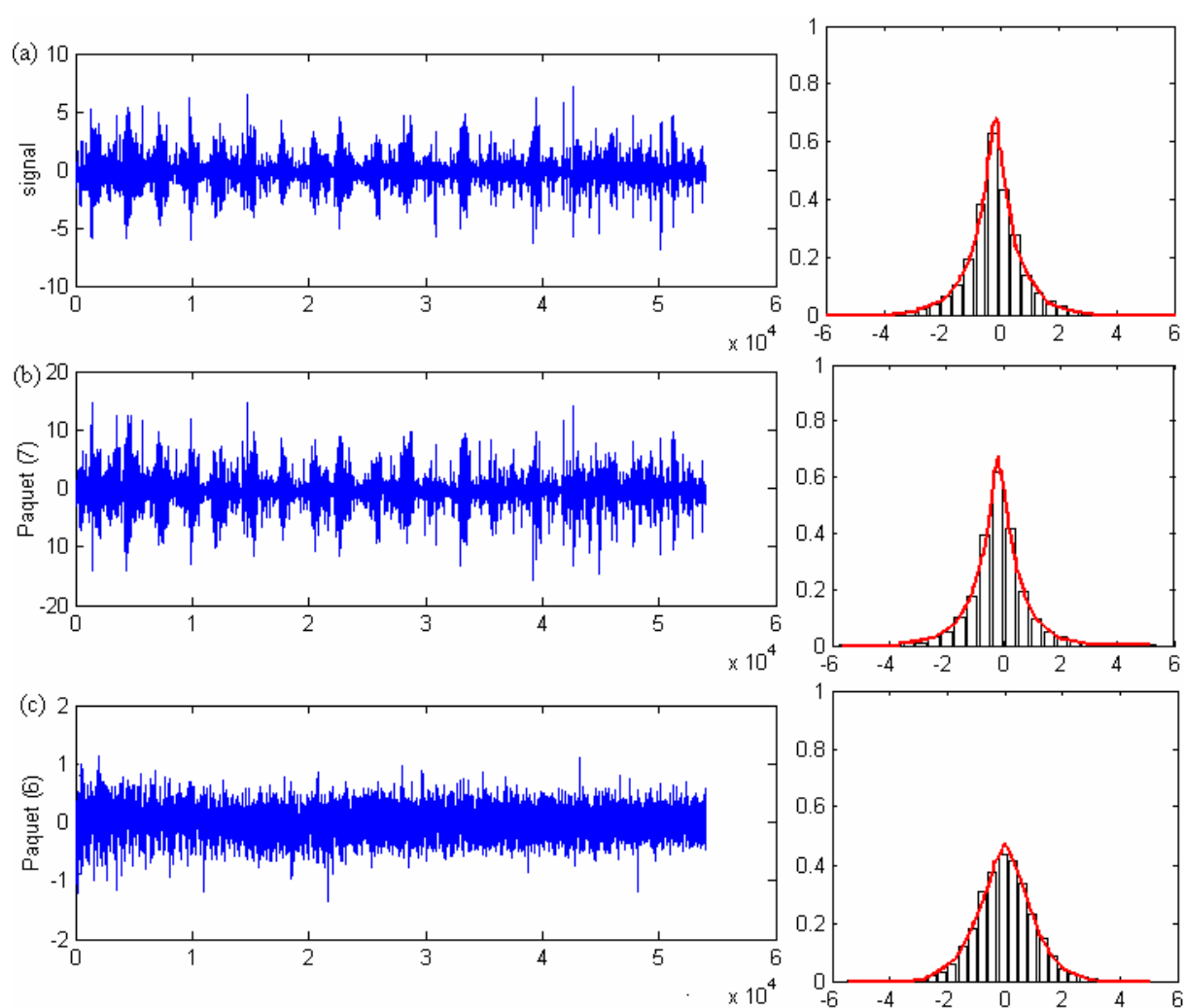


Figure 3.4. (a) Histogramme du signal femme tracé au dessus, ($\alpha = 0.6227, \beta = 0.9352$). (b) Histogramme des coefficients du paquet 7 du même signal EMG tracés avec la GGD. ($\alpha = 0.6142, \beta = 0.9257$). (c) Histogramme des coefficients du paquet 6 tracé avec la GGD ($\alpha = 1.3853, \beta = 1.9207$).

3.4 Estimation de la distance de Kullback Leibler

L'estimation de la distance de K-L est une partie cruciale pour l'utilisation du modèle statistique. Elle est basée sur le principe du rapport de vraisemblance [Do02]. Il a déjà été précisé que la distance de K-L entre deux densités de probabilités f_{θ_i} et f_{θ_j} d'une variable aléatoire X est définie par l'expression suivante :

$$D(f(X, \theta_i), f(X, \theta_j)) = \int_x f(x, \theta_i) \log \frac{f(x, \theta_i)}{f(x, \theta_j)} dx$$

Comme cela a été démontré dans le paragraphe précédent, les coefficients des paquets d'ondelettes des signaux EMG utérins suivent une loi gaussienne généralisée. La distance de K-L peut donc être complètement définie par la substitution de l'équation (3.2) dans l'équation (3.1). Après quelques étapes de calcul détaillées en annexe C, nous avons obtenu l'expression suivante de la distance de K-L:

$$D(f(., \alpha_i, \beta_i); f(., \alpha_j, \beta_j)) = \log \left(\frac{\beta_i \alpha_j \tau(1/\beta_j)}{\beta_j \alpha_i \tau(1/\beta_i)} \right) + \left(\frac{\alpha_i}{\alpha_j} \right)^{\beta_j} \frac{\tau((\beta_j + 1)/\beta_i)}{\tau(1/\beta_i)} - \frac{1}{\beta_i} \quad (3.7)$$

Cette expression se simplifie notablement lorsqu'on limite l'analyse aux signaux suivant la loi gaussienne ($\beta = 2$). On sait en particulier que cette hypothèse simplificatrice est parfaitement valide en l'absence d'événements significatifs dans un enregistrement. La distance de K-L devient alors:

$$D(f(., \alpha_i), f(., \alpha_j)) = K_{ij} = \log \left(\frac{\alpha_j}{\alpha_i} \right) + \frac{1}{2} \left(\frac{\alpha_i}{\alpha_j} \right)^2 - \frac{1}{2} \quad (3.8)$$

La distance de K-L est non symétrique. Pour résoudre ce problème, nous proposons d'utiliser :

$$K = K_{ij} + K_{ji} \quad (3.9)$$

$$K = \frac{1}{2} \left[\left(\frac{\alpha_i}{\alpha_j} \right)^2 + \left(\frac{\alpha_j}{\alpha_i} \right)^2 - 2 \right] \quad (3.10)$$

Si N est la longueur de la séquence numérique $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$, l'estimation du paramètre α est donnée par :

$$\hat{\alpha} = \frac{2}{N} \sqrt{\left(\sum_{i=1}^N |x_i|^2 \right)} \quad (3.11)$$

Alors l'estimation de K est:

$$\hat{K} = \frac{1}{2} \left[\left(\frac{\hat{\alpha}_i}{\hat{\alpha}_j} \right)^2 + \left(\frac{\hat{\alpha}_j}{\hat{\alpha}_i} \right)^2 - 2 \right] \quad (3.12)$$

La fonction définie ci-dessus est une fonction du rapport $\hat{\alpha}_i / \hat{\alpha}_j$.

Quand toutes les séquences \mathbf{x}_p suivent la même distribution gaussienne, $\alpha_i = \alpha_j = \alpha$, alors $K_{ij} = 0$.

Considérons le cas où deux séquences $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{iN})$ et $\mathbf{x}_j = (x_{j1}, x_{j2}, \dots, x_{jN})$ sont utilisées pour l'estimation de K_{ij} , et supposons dans un premier temps que les échantillons successifs x_{in} sont indépendants. $\hat{\alpha}_i$ et $\hat{\alpha}_j$ sont estimés en utilisant l'équation (3.11).

$N \left(\frac{\hat{\alpha}_i}{\alpha} \right)^2$ et $N \left(\frac{\hat{\alpha}_j}{\alpha} \right)^2$ suivent une loi du chi deux à N degrés de liberté. $z = \left(\frac{\hat{\alpha}_i}{\hat{\alpha}_j} \right)^2$ suit alors

une loi de Fisher $F_{N,N}$ avec $E(z) = \frac{N}{N-2}$ et $\text{var}(z) = \frac{4N(N-1)}{(N-2)^2(N-4)}$.

De la même manière, $\frac{1}{z} = \left(\frac{\hat{\alpha}_j}{\hat{\alpha}_i} \right)^2$ suit la loi de Fisher avec la même espérance et la même

variance.

Retournons à la distance de K-L estimée :

$$\hat{K} = \frac{1}{2} \left[\left(\frac{\hat{\alpha}_i}{\hat{\alpha}_j} \right)^2 + \left(\frac{\hat{\alpha}_j}{\hat{\alpha}_i} \right)^2 - 2 \right] \quad (3.13)$$

L'espérance de la distance de K-L estimée \hat{K} devient :

$$E(\hat{K}_{ij} + \hat{K}_{ji}) = E(\hat{K}) = \frac{2}{N-2} \quad (3.14)$$

alors que

$$E(\hat{K}_{ij}) = \frac{1}{N-2} \quad (3.15)$$

3.5 Estimation de la distribution de \hat{K}

Le paragraphe précédent montre que l'estimation de la distance de Kullback Leibler \hat{K} entre deux séquences \mathbf{x}_i et \mathbf{x}_j peut être obtenue à partir de l'estimation de la variance correspondant à ces intervalles, dans le cas gaussien, c'est-à-dire dans le cas où les deux segments ne contiennent que du bruit.

L'approche que nous avons utilisée consiste alors à estimer tout d'abord la distribution de \hat{K} à partir de M séquences de bruit gaussien de longueur N, puis à modéliser la distribution ainsi obtenue.

Dans un deuxième temps, les distributions de \hat{K} obtenues sur chaque nœud de la décomposition de la base de signaux sur l'arbre des paquets d'ondelettes seront comparées à cette distribution de référence. Les nœuds produisant des distributions très éloignées seront considérés comme les plus représentatifs de la présence d'événements.

La figure 3.5 illustre ce principe : un signal EMG bruité contenant un certain nombre d'événements (ici quatre) est projeté sur chacun des nœuds de l'arbre de décomposition en paquets d'ondelettes. Sur la figure sont indiquées les séquences de coefficients obtenues pour les nœuds 6 et 7, le premier ne contenant que du bruit alors que le second met bien en évidence les événements EMG. Après découpage en intervalles de longueur N (constante), chaque paire d'intervalles produit une valeur de \hat{K} . A chaque nœud correspond donc une distribution. Seuls les nœuds ne contenant que du bruit sont censés présenter une distribution correspondant à l'équation (3.13).

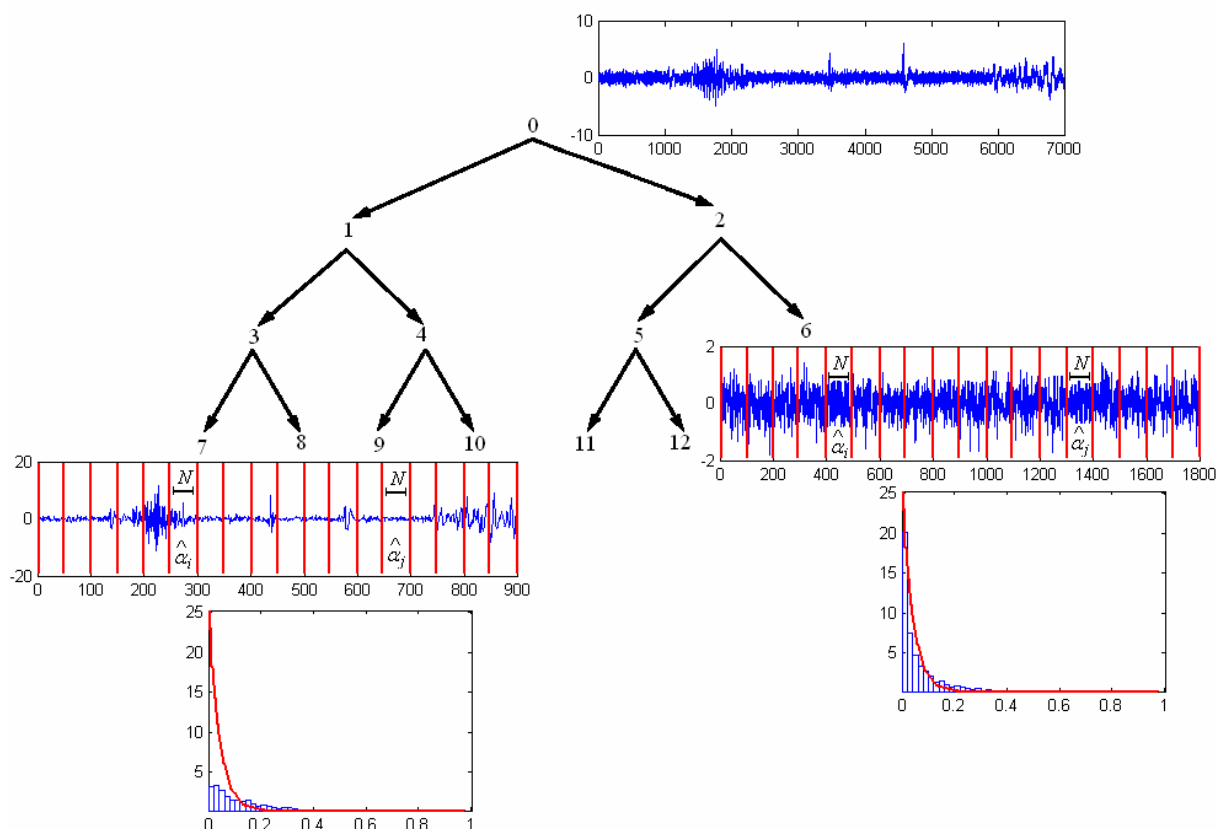


Figure 3.5. Découpage des coefficients des paquets d'ondelettes en fenêtres de longueurs égales pour l'estimation de distance de K-L avec les histogrammes associés aux paquets 6 et 7..

3.5.1 Distribution obtenue dans le cas où il n'y a pas d'événements dans les enregistrements ($\alpha_i = \alpha_j$)

L'étude de la distribution de \hat{K} dans le cas où il n'y a pas d'événements ($\alpha_i = \alpha_j$) servira de base pour décider si un paquet contient des événements ou pas, et par conséquent d'éliminer ce paquet lors de l'application des algorithmes de détection. Comme cela a été indiqué au paragraphe précédent, l'idée consiste à tracer l'histogramme de \hat{K} à partir d'une séquence de bruit blanc gaussien et d'étudier ses caractéristiques statistiques. Commençons par le tracé de l'histogramme de \hat{K} obtenu par simulation. La simulation a été réalisée en générant 500 segments de bruit blanc gaussien $N(0, 1)$, chacun ayant une longueur égale à 1000 (nombre de points). Au total la longueur du signal utilisé est de 500.000 points. La longueur de la fenêtre utilisée dans ce cas est $N = 100$. 12,497,500 valeurs de \hat{K} ont été estimées. La figure 3.6

montre l'histogramme de \hat{K} ainsi obtenu. \hat{K} est calculé en utilisant l'équation (3.13) après estimation de $\hat{\alpha}_i$ et $\hat{\alpha}_j$.

Puisque nous ne disposons pas d'une expression analytique de la distribution de \hat{K} , l'idée est d'approcher l'histogramme de \hat{K} par une distribution connue ayant au moins la même forme générale et la même espérance. L'histogramme empirique de \hat{K} (figure 3.6), construit par simulation n'est fonction que d'un seul paramètre N . N représente le degré de liberté, correspondant au nombre des points considérés dans chaque segment.

Nous avons donc cherché à modéliser l'histogramme par une distribution inspirée par la forme générale de l'histogramme et ne dépendant que d'un seul paramètre.

Tenant en compte la forme de l'histogramme tel que présenté sur la figure 3.6, nous avons choisi de prendre comme première approximation de cet histogramme la loi exponentielle, dépendant elle-même d'un seul paramètre λ . Sa densité de probabilité est définie par :

$$f(x) = \lambda.e^{-\lambda.x} \text{ avec } E(x) = \frac{1}{\lambda}. \quad (3.16)$$

Un ajustement entre $f(x)$ et l'histogramme est facilement obtenu en égalant les deux espérances :

$$\frac{1}{\lambda} = \frac{2}{N-2} \quad (3.17)$$

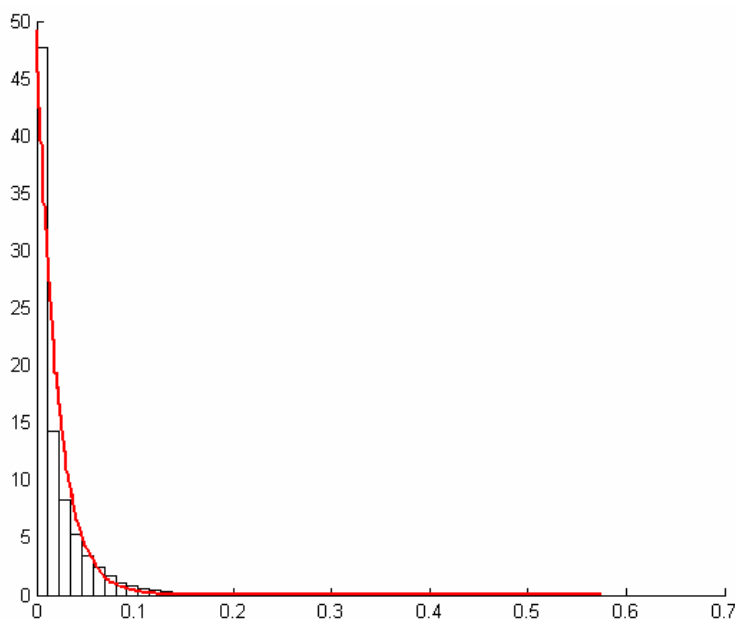


Figure 3.6. Histogramme de \hat{K} associé à la distribution exponentielle d'espérance $\frac{2}{N-2}$.

Pour ce qui concerne la modélisation à partir de bruit réel, les segments de bruit de l'ensemble d'apprentissage de la base REEL1 ont été concaténés. Un segment de 45549 points est alors formé. Ce segment est décomposé en intervalles égaux de 100 points ($N=100$). \hat{K} est calculé entre chaque paire d'intervalles après estimation de $\hat{\alpha}_i$ et $\hat{\alpha}_j$. L'histogramme de \hat{K} est présenté à la figure 3.7 avec la distribution exponentielle d'espérance $\frac{2}{N-2}$. On voit bien que l'histogramme de \hat{K} suit approximativement la distribution exponentielle, ce qui montre entre autres que l'hypothèse d'indépendance des échantillons successifs n'est pas critique.

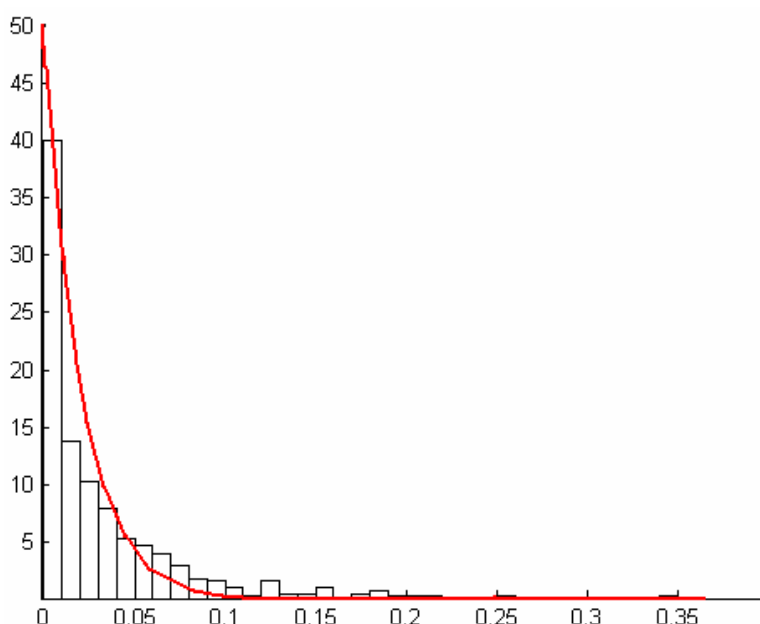


Figure 3.7. Histogramme de \hat{K} des segments concaténés du bruit réel (GIREEL1) associé à la distribution exponentielle d'espérance $\frac{2}{N-2}$.

3.5.2 Cas où des événements sont présents dans les enregistrements ($\alpha_i \neq \alpha_j$)

Ce cas permet d'évaluer le comportement de la distance de K-L estimée \hat{K} quand des événements sont présents dans un paquet donné. La figure 3.8 montre l'histogramme du signal femme présenté à la figure 3.4. Il est évident que l'histogramme de \hat{K} est très loin de la distribution attendue en l'absence des événements.

Ceci permet de dire que des événements sont bien mis en évidence dans les paquets où l'histogramme de \hat{K} est loin de la loi exponentielle, alors qu'il n'y a pas d'événements identifiables dans les paquets où l'histogramme de \hat{K} suit approximativement une loi exponentielle.

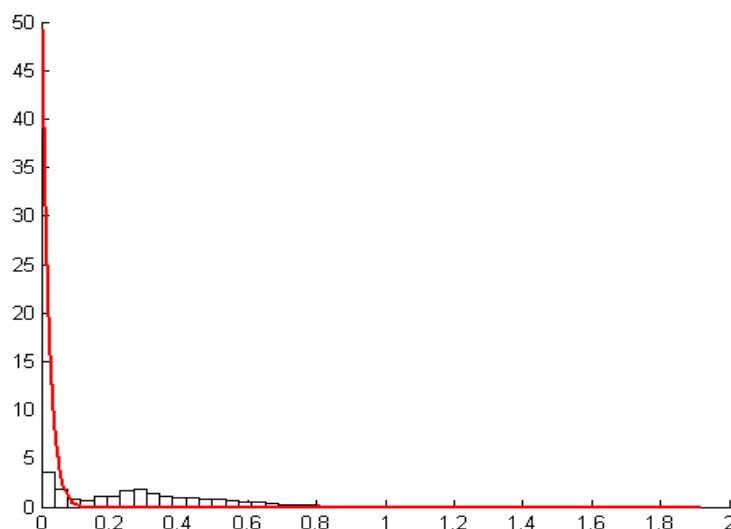


Figure 3.8. Histogramme de \hat{K} associé à la distribution exponentielle d'espérance $\frac{2}{N-2}$, lorsqu'il existe des événements dans les enregistrements.

3.6 Sélection des paquets d'ondelettes

3.6.1 Caractérisation des paquets d'ondelettes pour la détection

La séquence des coefficients de chaque paquet d'ondelettes est découpée en intervalles égaux de longueur N et la distance de K-L est estimée entre chaque paire d'intervalles jusqu'à balayer toutes les paires possibles. Les paramètres α_i et α_j sont estimés pour chaque paire d'intervalles pour permettre le calcul de la distance de K-L estimée.

L'objectif de cette étude est de ne retenir que les paquets qui extraient efficacement les événements du bruit. La caractérisation d'un paquet se fait donc de la façon suivante : s'il n'y a que du bruit, ou si la différenciation entre événements et bruit se fait dans de mauvaises conditions pour un paquet donné, alors la distribution de \hat{K} suit une loi proche de la loi exponentielle d'espérance $\frac{2}{N-2}$.

La comparaison entre la distribution approchée de \hat{K} et la distribution exponentielle est faite en utilisant la statistique de Kolmogorov Smirnov (K-S) [Sap90]. Cette statistique du test de K-S D_{\max} , résulte du calcul de la différence maximale entre les deux fonctions de répartition des lois expérimentale et théorique.

En résumé, nous définissons les deux hypothèses suivantes :

H_0 : \hat{K} suit une loi exponentielle (pas d'événements détectables dans le nœud correspondant).

H_1 : \hat{K} ne suit pas une loi exponentielle (des événements sont détectés) .

Il peut ensuite y avoir deux approches dans l'utilisation de la statistique D_{\max} . La première consiste à définir un seuil pour D_{\max} à partir duquel on considère qu'il y a des changements détectés dans les paquets définis. La deuxième approche consiste à faire un tri par ordre décroissant des valeurs de D_{\max} calculées sur chaque nœud de l'arbre de décomposition, et de sélectionner ensuite un nombre donné de paquets correspondant aux plus grandes valeurs de D_{\max} .

3.6.2 Effet de la longueur de la fenêtre N

La loi exponentielle n'est pas un modèle fin pour la distribution de \hat{K} . Nous avons voulu vérifier que la valeur de N n'avait pas d'influence sur la différence entre le modèle et l'histogramme estimé. Nous nous sommes d'abord placés en situation de simulation par un bruit blanc gaussien, puis en situation plus réaliste à partir du bruit réel d'enregistrement pour lequel l'indépendance des échantillons successifs des séquences de coefficients n'est plus garantie.

En simulation, quatre signaux sont utilisés pour calculer \hat{K} , chaque signal étant un bruit blanc gaussien de moyenne nulle et de variance égale à 1. Chaque signal est décomposé en 448 fenêtres de longueur N . La longueur totale de chaque signal est alors $448*N$. nous obtenons 100.000 valeurs de \hat{K} pour lesquelles on trace les histogrammes. Quatre valeurs de N (50, 100, 200, 500) ont été testées. La loi exponentielle correspondante a été prise pour $\frac{1}{\lambda} = \frac{2}{N-2}$. La valeur de D_{\max} a été calculée pour les 4 valeurs de N . La même procédure a été appliquée sur des signaux de bruit réel issus de l'ensemble d'apprentissage G1REEL1 de

la base REEL1 décrite à l'annexe A. Le tableau 3.4 montre les valeurs de D_{max} calculées entre les distributions exponentielle et expérimentale de distance de K-L estimée \hat{K} pour 4 valeurs de N ($N = 50$, $N = 100$, $N = 200$ et $N = 500$).

	N = 50	N = 100	N = 200	N = 500
Signaux Simulés	0.1465	0.1545	0.1546	0.1615
Bruit réel	0.1550	0.1504	0.1562	0.1589

Tableau 3.4. D_{max} calculée pour différentes valeurs de N dans le cas des signaux simulés et dans le cas du bruit réel.

En utilisant les signaux synthétiques, 20 valeurs de D_{max} ont été calculées pour $N = 50$. Nous avons obtenu $\bar{D}_{max} = 0.1550$ et $\sigma_{D_{max}} = 0.0094$. Au vu de ces résultats, nous pouvons considérer que 50 points est parfaitement valide, puisqu'une faible différence est constatée lorsque des fenêtres beaucoup plus grandes sont utilisées.

3.6.3 Construction de l'algorithme de sélection de la meilleure base

L'étape précédente de caractérisation des paquets d'ondelettes identifie les nœuds ayant la meilleure capacité à mettre en évidence des événements d'EMG. Cependant, comme l'arbre contient des paquets redondants, l'étape suivante consiste à minimiser cette redondance.

Pour définir l'algorithme de sélection de la meilleure base, nous avons choisi de suivre approximativement la même démarche que celle proposée par Hitti et Lucas [Hit98].

Définissons un nœud père (NP) comme étant une connexion entre deux branches dont les extrémités sont appelées les nœuds fils (NF). Un nœud père au niveau j est le fils du nœud auquel il est rattaché au niveau $j-1$. Un composant est l'association d'un NP et de ses deux fils (NF).

Les étapes de notre algorithme de sélection de la meilleure base sont alors définies comme suit:

- a. Sélection des paquets significatifs:** À partir de l'arbre binaire des paquets d'ondelettes, on construit un arbre de même structure, dans lequel les nœuds prennent les valeurs "1" ou "0" selon les valeurs D_{max} résultant du test de Kolmogorov Smirnov. Comme indiqué

précédemment, deux méthodes de sélection peuvent être utilisées, la première est de se référer à un seuil s pour ne sélectionner que des paquets pour lesquels $D_{max} > s$. La deuxième est de déterminer un nombre P des paquets *a priori* et de choisir alors ceux qui possèdent les plus grandes valeurs de D_{max} (figure 3.9.a).

$$\left\{ \begin{array}{l} -1 \text{ si } D_{max} > s \text{ (méthode 1)} \\ \text{ou} \\ \text{si } D_{max} \text{ est parmi les } P \text{ plus grandes valeurs (méthode 2)} \\ -0 \text{ si non} \end{array} \right.$$

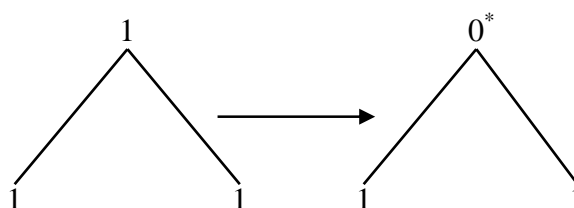
b. Réduction du nombre de nœuds pour la meilleure base (les règles correspondantes sont illustrées sur la figure 3.9.b) :

Pour cette phase de réduction, on utilise trois valeurs possibles pour les nœuds :

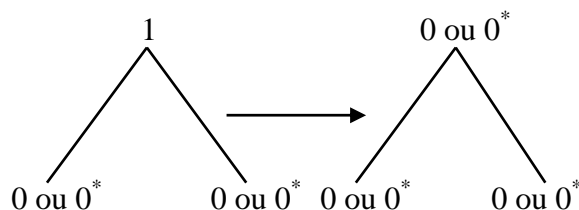
- 1 : le nœud contient bien de l'information originale sur les événements.
- 0 : le nœud ne contient pas d'information sur les événements.
- 0* : le nœud contient de l'information, mais celle-ci a déjà été identifiée sur un autre nœud.

Les règles de réduction proposées sont alors les suivantes :

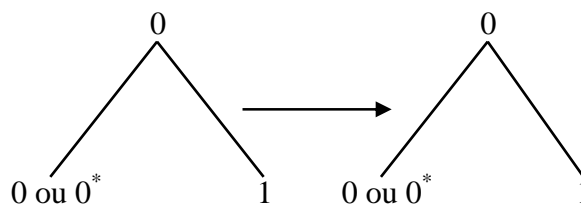
- si tous les nœuds dans un composant possèdent la valeur "1", mettre NP à 0* car toute l'information est contenue dans les nœuds fils NF :



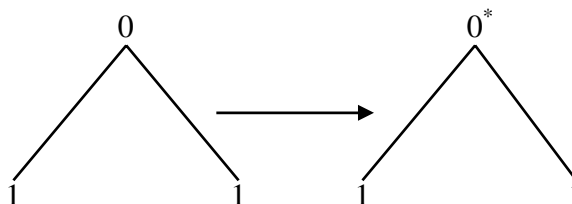
- si NP = 1 et les deux NF = 0 ou 0*, mettre NP à "0" car l'information est déjà prise en considération par les nœuds fils dans le cas où un NF = 0*, (le cas où les deux NF = 0 est par nature extrêmement rare) :



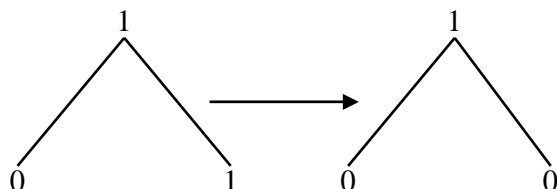
- si $NP = 0$ et l'un des deux fils $NF = 1$, laisser le NF correspondant à 1 car l'information se trouve bien dans ce noeud :



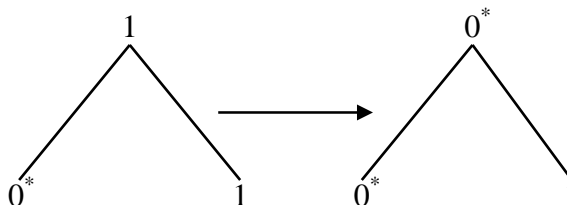
- si $NP = 0$ et les deux fils $NF = 1$, alors $NP = 0^*$ car NP contient probablement des informations mal identifiées à ce niveau de décomposition alors qu'elles sont plus clairement mises en évidence dans les NF :



- si $NP = 1$ avec $NF1 = 0$ et $NF2 = 1$, mettre alors $NF = 0^*$ car dans ce cas NP et NF contiennent les mêmes informations :



- si $NP = 1$ avec $NF1 = 0^*$ et $NF2 = 1$, laisser alors $NF2 = 1$ et mettre $NP = 0^*$, car toute l'information est incluse, implicitement ou explicitement, dans les NF :



c. Sélection finale de la meilleure base.

Elle correspond à ne conserver que les nœuds ayant la valeur finale 1 (figure 3.9.c).

Cette dernière étape correspond à réduire la redondance en ne conservant pas deux paquets jouant le même rôle en matière de détection. Comme le fait de conserver à chaque niveau le père ou ses fils n'a pas été garanti, la sélection finale ne constitue pas une base complète.

Cette réserve ne pose cependant aucun problème dans notre cas puisque l'objectif est de retenir les paquets significatifs en terme de présence de ruptures puis d'appliquer un algorithme de détection sur les paquets retenus. En aucun cas, il ne sera nécessaire ensuite de revenir à une approximation du signal initial par reconstruction.

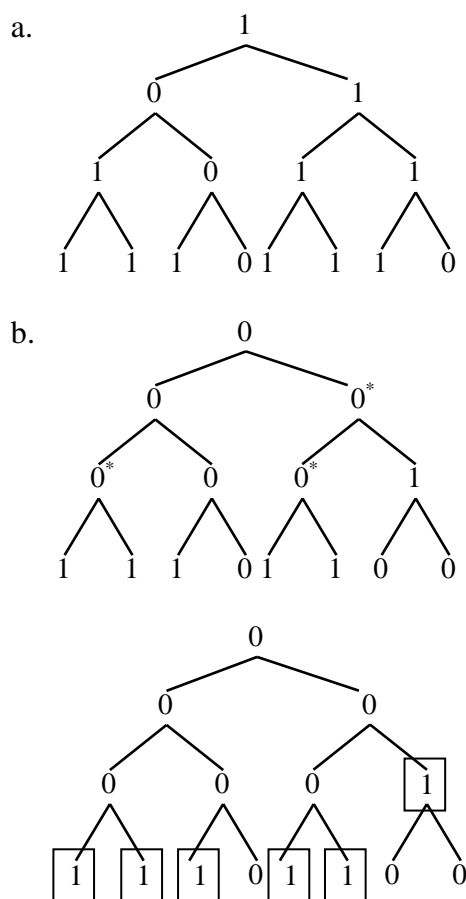


Figure 3.9. Etapes de construction de la meilleure base du signal : a- Mettre à "1" les paquets ayant les plus grandes valeurs de D_{\max} et à "0" les autres. b- Modification des valeurs des nœuds. c- La meilleure base constituée des nœuds à 1.

3.6.4 Paramètres attachés à l'algorithme de sélection de la meilleure base

Les performances de cet algorithme vont dépendre principalement :

- du choix de l'ondelette et donc de la capacité de ses filtres à concentrer l'information utile d'un événement au sein d'un seul paquet. Il s'agit de choisir une ondelette aussi sélective que possible. L'algorithme sera donc testé sur plusieurs types d'ondelettes et l'ondelette retenue sera celle qui produira les meilleurs performances en matière de détection.

- de la pertinence du choix du seuil de la statistique de Kolmogoro v Smirnov D_{\max} ou du nombre de paquets ayant les plus grandes valeurs de D_{\max} . Pour choisir un seuil, les valeurs de D_{\max} seront triées par ordre décroissant, puis la valeur du seuil sera déterminée en examinant la suite des différences des valeurs de D_{\max} .

3.7 Application de la méthode de sélection de la meilleure base

3.7.1 Résultats de la méthode appliquée sur les signaux synthétiques

Comme le contenu fréquentiel de l'EMG utérin est un des éléments majeurs pour la discrimination entre le bruit et les principaux événements contenus dans le signal [Kha99b], l'algorithme de sélection de la meilleure base est tout d'abord appliqué sur un ensemble d'apprentissage composé de signaux simulés dans lesquels les changements sont purement fréquentiels. Les signaux utilisés appartiennent à l'ensemble d'apprentissage G1SIM1 de la base SIM1 décrite à l'annexe A. Rappelons que chacun de ces signaux est composé de 4 segments correspondant à différentes bandes fréquentielles dont les bornes sont définies par référence à la moitié de la fréquence d'échantillonnage $F_e / 2$ ([0.01 ; 0.4], [0.01 ; 0.6], [0.01 ; 0.8], et [0.01 ; 0.999] Hz). Le choix de ces bandes fréquentielles simule la superposition fréquentielle qui existe dans les signaux EMG utérin. Chaque segment est composé de 2000 échantillons et chaque signal contient 3 instants de changement : 2000, 4000, et 6000. La longueur totale de chaque signal est de 8000 points. 1000 signaux ont été simulés.

La décomposition en paquets d'ondelettes a été faite jusqu'au niveau 3 en utilisant l'ondelette Symlet 5.. Pour décider quels sont les paquets capables de refléter la présence des ruptures, moyenne et écart-type des valeurs statistiques du test de K-S, D_{\max} , ont été calculés pour chaque paquet (il y a 14 paquets dans ce cas). La figure 3.10 montre les valeurs de D_{\max} dans l'ordre décroissant. Dans l'arbre de décomposition, la valeur "1" est donnée aux paquets ayant les plus grandes valeurs de D_{\max} .

En examinant la figure 3.10, on peut voir qu'il existe un seuil évident pour D_{\max} égal à 0.5. La valeur "1" est alors donnée aux paquets 11, 5, 12, 14, 2 et 6 (figure 3.11.a).

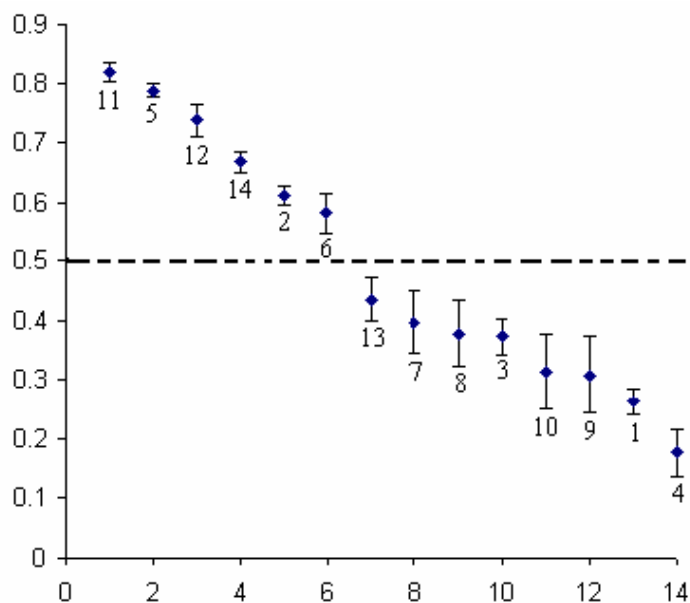


Figure 3.10. Moyennes et écarts-types des valeurs de la statistique D_{max} de K-S obtenues pour les 14 paquets à partir de 1000 signaux simulés. Axe X : unité arbitraire. Axe Y: Valeurs de D_{max} .

Après l'application de l'algorithme de sélection de la meilleure base, les paquets 6, 11, et 12 ont été retenus comme étant la meilleure base "incomplète" pour les signaux synthétiques décrits dans ce paragraphe (figure 3.11.b).

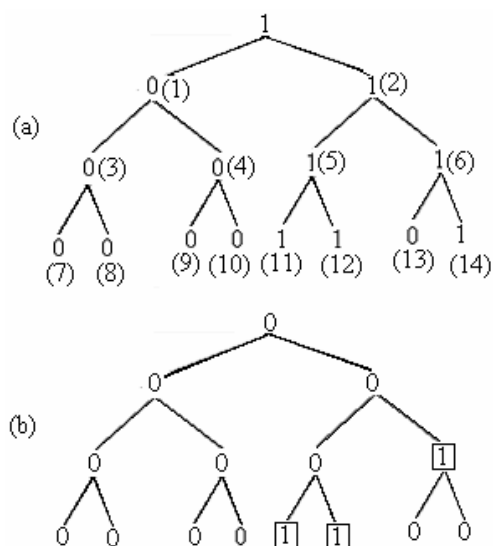


Figure 3.11. (a) Classification de chacun des paquets selon les valeurs de D_{max} (le numéro de chaque paquet est indiqué à l'intérieur des parenthèses). (b) Sélection de la meilleure base (incomplète dans ce cas).

La figure 3.12 représente un signal de la base SIM1, avec les tracés des coefficients des paquets 6, 11 et 12. On voit bien que ces paquets identifient les quatre segments du signal, ce

qui vérifie que l'application de l'algorithme du choix de la meilleure base a conservé les informations utiles.

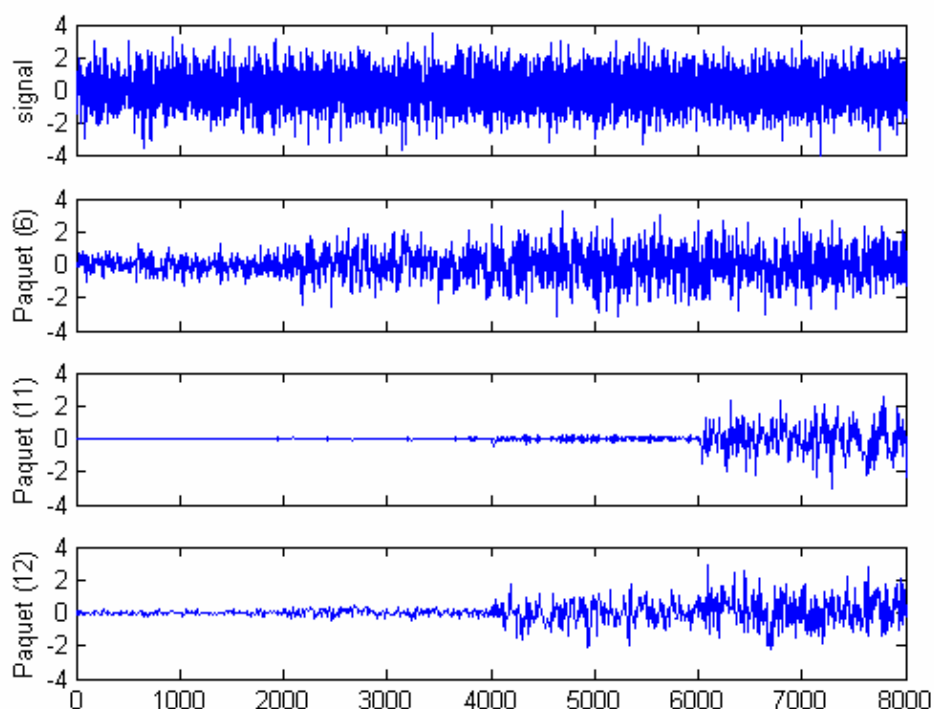


Figure 3.12. Décomposition d'un signal de la base SIMI en paquets d'ondelettes. Les coefficients des paquets retenus 6, 11 et 12 sont représentés. L'ondelette utilisée est Symlet no.5. Axes des abscisses : nombre des points. Axes des ordonnées : unités arbitraires.

Le paragraphe 4.5 reviendra plus en détail sur la pertinence de la méthode de réduction lorsqu'un algorithme de détection aura été défini et qu'une mesure de performance pourra alors être faite à partir des taux de bonnes détections et de fausses alarmes.

3.7.2 Résultats de la méthode appliquée sur les signaux réels

Dans le cadre de ce chapitre, nous utilisons l'ensemble d'apprentissage G1SIMREEL constitué des 100 signaux de la base SIMREEL décrite à l'annexe A. Rappelons que chaque signal est composé de 10 événements de différents types appartenant à différents femmes et termes (contractions, ondes d'Alvarez, ondes LDBF, mouvements actifs fœtaux...) identifiés par un expert, et que ces événements ont été chaînés et contaminés par du bruit blanc (SNR = 10 db). Les enregistrements réels complets seront utilisés lors de la validation de la procédure globale de segmentation et de classification. Les signaux ont été projetés sur l'ensemble des nœuds de l'arbre de décomposition en paquets d'ondelettes, l'ondelette Symlet 5 est aussi

utilisée. Après application de la première étape de l'algorithme de sélection de la meilleure base, seuls les paquets 1, 3, 7 et 8 ont été sélectionnés (figure 3.13). Ces paquets ont les bandes fréquentielles [0-4], [0-2], [0-1] et [1-2] Hz respectivement. Ils ont été sélectionnés grâce à la statistique de K-S D_{\max} en mettant le seuil à 0.25. Le choix de cette valeur de seuil était facile car la distinction entre les paquets à grandes valeurs de D_{\max} et les autres est claire (figure 3.13).

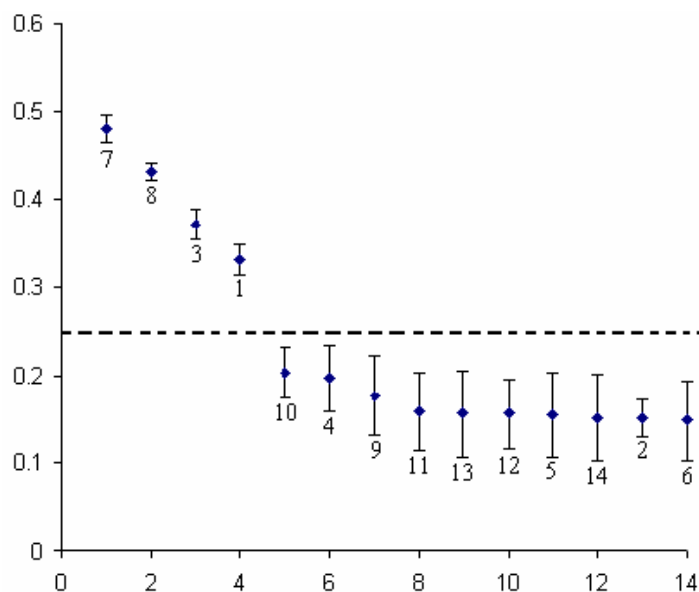


Figure 3.13. Moyennes et écarts-types des valeurs de la statistique D_{\max} de K-S obtenues pour les 14 paquets à partir de 100 signaux d'EMG utérin de la base REELSIM. Axe X : unité arbitraire. Axe Y: Valeurs de D_{\max} .

La valeur "1" est alors donnée aux paquets 1, 4, 7, et 8 (figure 3.14.a). Après l'application de l'algorithme de sélection de la meilleure base, seuls les paquets 7 et 8 ont été finalement retenus (figure 3.14.b).

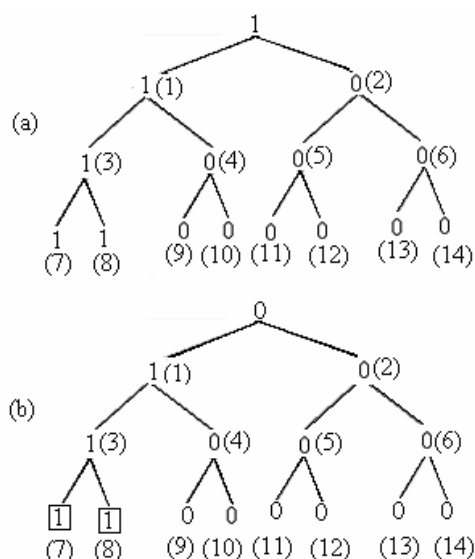


Figure 3.14. Sélection de la meilleure base dans le cas de l'EMG utérin.

La figure 3.15 représente un signal de la base SIMREEL décrite à l'annexe A, avec les tracés des coefficients des paquets 7 et 8.

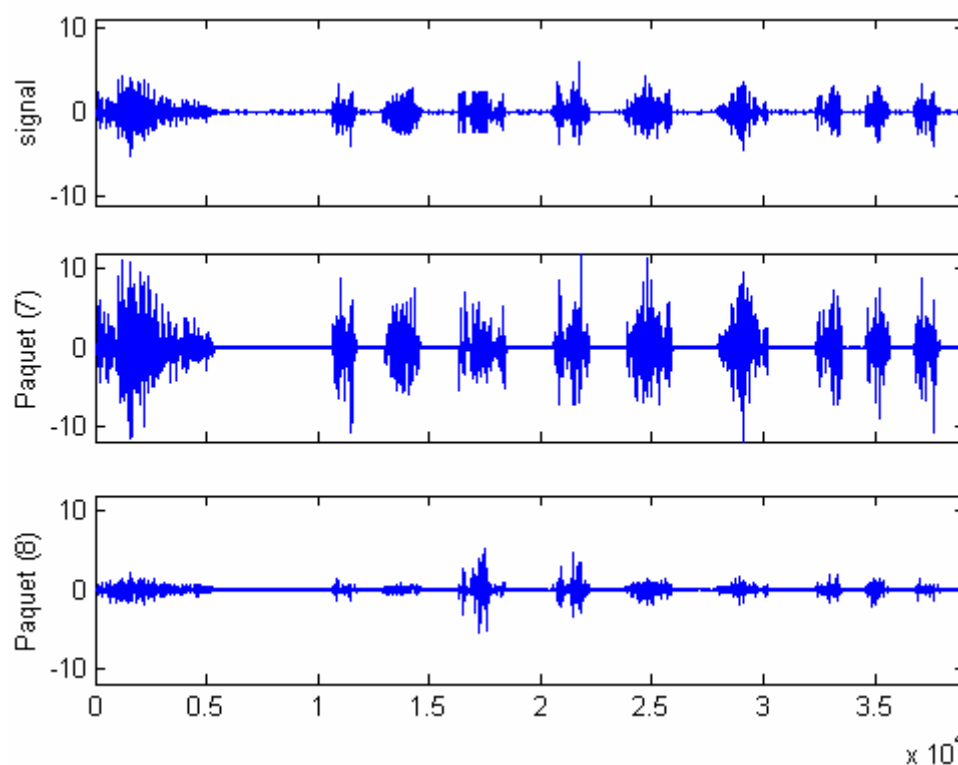


Figure 3.15. Décomposition d'un signal de la base SIMREEL en paquets d'ondelettes. Les coefficients des paquets retenus 7 et 8 sont représentés. L'ondelette utilisée est daubechies no.3. Axes des abscisses : nombre des points. Axes des ordonnées : Unités arbitraires.

On voit bien que ces paquets identifient les événements contenus dans le signal, ce qui illustre le fait que l'application de l'algorithme du choix de la meilleure base a garanti la conservation

des informations utiles. Là encore, une évaluation plus rigoureuse sera proposée au paragraphe 4.5.

3.8 Discussion et Conclusion

La sélection des nœuds pertinents de l'arbre obtenu par décomposition en paquets d'ondelettes est généralement guidée par un critère défini selon les objectifs de l'application. Ce critère est choisi au vu de la connaissance des propriétés des données (souvent modélisées par des distributions statistiques) ou de la disponibilité d'un ensemble d'apprentissage. L'entropie est souvent utilisée quand l'objectif correspond à une perte minimale d'information. L'information (ou la distance) de Kullback Leibler est bien adaptée aux problèmes de détection et de classification quand des rapports de vraisemblance peuvent être calculés. La sélection finale des paquets présentant les changements entre les événements, après l'application de l'algorithme du choix de la meilleure base, aboutit à un choix pertinent des paquets significatifs en terme de détection d'événements. La distance de Kullback Leibler, de ce fait, est un bon outil statistique pour décider de la capacité d'un paquet à détecter les ruptures dans une classe de signaux donnée.

Nous avons montré que les activités contenues dans l'EMG utérin suivaient une distribution gaussienne généralisée. Quant au bruit d'enregistrement (extrait entre les activités par sélection manuelle), il suit une loi gaussienne. Par conséquent, en comparant des segments décomposés en paquets d'ondelettes contenant seulement du bruit, la distance de K-L entre ces segments doit être théoriquement nulle.

Comme nous pouvons seulement avoir accès à une estimation de cette distance, nous avons défini par simulation un histogramme de la distance K-L estimé sous l'hypothèse qu'il n'y a aucun changement dans le signal. Les meilleurs paquets étant ceux qui produisent la différence la plus élevée entre les événements et le bruit, la distribution de la distance de K-L estimée pour ces paquets doit être loin de la distribution estimée pour les paquets sans aucun changement. La statistique de Kolmogorov-Smirnov D_{max} a été employée pour évaluer la distance entre les distributions. Nous avons choisi de modéliser l'histogramme de la distance de K-L par une distribution exponentielle, de telle sorte que l'ajustement entre les histogrammes de distance de K-L estimée et de la loi exponentielle soit limité à l'égalité de leurs espérances.

Nous n'avons pas proposé de méthode unique en terme de sélection des meilleurs paquets, cette sélection pouvant se faire soit par un seuil sur D_{max} , soit en limitant le nombre de paquets sélectionnés.

Les résultats montrés sur les figures 3.10 (signaux synthétiques) et 3.13 (signaux réels) ont illustrés une démarche possible pour la détermination du seuil de sélection des paquets en considérant le pas de décroissance de D_{max} .

De plus il a été demandé à un expert de décider quels sont les paquets qui représentent au mieux l'EMG utérin. Soulignons à ce niveau que le nombre d'experts capables d'interpréter un enregistrement d'EMG utérin est encore très limité aujourd'hui, et que le niveau de concordance entre experts n'est pas de 100%. Néanmoins l'avis de l'expert sur la sélection des meilleurs paquets a confirmé le choix effectué grâce à l'application de la méthode de la sélection de la meilleure base.

L'objectif du chapitre suivant est d'appliquer un algorithme de détection séquentiel sur les paquets choisis afin de pouvoir quantifier les performances de l'algorithme de choix de la meilleure base.

Chapitre 4

Détection et segmentation de l'EMG utérin

4.1 Introduction

La segmentation est un problème primordial en traitement du signal. C'est une approche classique pour le traitement des signaux présentant des non-stationnarités rapides (des ruptures). Elle constitue une première étape importante de traitement en vue de la reconnaissance des formes et du diagnostic. Une non-stationnarité rapide ou rupture ou encore une transition sera considérée ici comme une phase transitoire de courte durée par rapport à la fenêtre d'observation. La détection d'une rupture sera quant à elle la version élémentaire de la segmentation d'un signal.

Cette étape de détection est souvent abordée sous l'angle statistique du test d'hypothèses [Ant01], [Bri92], [Kha99], [Rav98]. Dans la littérature [Bri92], [Rav98], [Wen96], on trouve plusieurs méthodes de détection dans le domaine temporel, comme les détecteurs de saut appliqués à certains paramètres du signal : saut de moyenne, de variance, voire des moments d'ordre supérieur pour la localisation des transitoires non gaussiens.

Une autre famille de détecteurs s'applique après modélisation du signal : le bruit additif est généralement supposé blanc, gaussien et stationnaire, alors que les événements sont modélisés comme des processus auto-régressifs. La segmentation est alors réalisée par une détection de changement de modèle. Un passage en revue rapide de la bibliographie montre que la

modélisation auto-régressive (AR) est souvent utilisée dans le domaine biomédical pour la segmentation de l'électromyogramme (EMG) utérin [Kha99], de l'électroencéphalogramme [Wen96], mais aussi dans le diagnostic industriel [Bri92], l'acoustique sous-marine [Rav98] ou la segmentation de la parole [Tho02].

La modélisation AR n'est évidemment pas la seule approche dans l'estimation du signal. Dans l'introduction de [Ant01], *Antoniadis et al.* énumèrent plusieurs méthodes de régression non paramétriques auxquelles se rajoutent progressivement, après 1990, les estimateurs non linéaires par seuillage des décompositions en ondelettes, qui donnent souvent de très bons résultats. La décomposition en ondelettes n'est pas, à proprement parler, une modélisation du signal, mais la transformation de celui-ci dans un autre domaine (temps-échelle). Cette décomposition permet l'étude détaillée de ses caractéristiques : par leur construction, les ondelettes sont des fonctions localisées en temps et en fréquence, elles s'apparentent donc aux représentations temps-fréquence. La modélisation du signal par décomposition en ondelettes est donc une méthode qui peut répondre aux exigences en matière de détection de signaux présentant des ruptures fréquentielles.

Dans notre cas, le principe retenu pour la détection est l'application d'algorithmes séquentiels sur les coefficients (et non sur les signaux reconstruits) des niveaux d'échelles (dans le cas de décomposition discrète en ondelettes) ou sur les coefficients des meilleurs paquets (dans le cas de décomposition en paquets d'ondelettes) préalablement choisis. La détection à partir des coefficients repose sur la recherche de la coïncidence en temps, à travers plusieurs échelles, de coefficients d'ondelettes significatifs. La décomposition dyadique peut cependant mener à un décalage important dans la localisation des changements. Le manque de synchronisme des détections au niveau de la séquence des coefficients des différents paquets ou niveaux d'échelles nous conduit à redéfinir la valeur vraie de l'instant de détection par rapport au signal original.

Dans la première phase de ce chapitre, nous allons étudier la théorie du décalage des coefficients d'ondelettes sur chaque niveau d'échelle et sur chaque paquet d'ondelettes. La correction des instants de détection sera ensuite étudiée pour recalibrer les instants de changement à leur vraie position. Chaque instant de détection pour un niveau d'échelle ou pour un paquet correspondra alors à un intervalle de détection au niveau du signal de départ. Comme la détection se fait sur plusieurs paquets ou niveaux d'échelles, un algorithme de fusion des intervalles de détection sera appliqué pour aboutir, lorsque cela sera justifié, à un seul instant de détection pour le signal original.

De plus, nous nous sommes attachés au développement d'un algorithme pouvant fonctionner en temps réel (§4.7).

Enfin, la performance des méthodes sera étudiée en les appliquant d'une part sur des signaux simulés, d'autre part sur la base de données disponibles des signaux réels d'EMG utérin.

4.2 Utilisation des coefficients d'ondelettes pour la détection

L'emploi direct des coefficients d'ondelettes sans recours à la reconstruction des signaux à partir de ces coefficients se révèle un outil de détection efficace, alliant de bonnes performances à un faible coût de calcul. En effet, si M est le nombre de points du signal, alors l'algorithme de décomposition discrète d'analyse et de reconstruction vaut $2.O(M)$, tandis qu'en utilisant seulement l'algorithme d'analyse sur les coefficients, le coût de calcul est réduit de moitié [Che05a].

De plus, l'un de nos buts est la détection en temps réel, c'est-à-dire effectuer la détection en même temps que l'acquisition des signaux, en ne tolérant que le retard indispensable à l'estimation et celui imposé par la détection. Dans le cadre de la décomposition en ondelettes ou en paquets d'ondelettes, à chaque acquisition d'un point du signal, il est tout à fait possible d'effectuer le calcul de nouveaux coefficients sur les différentes échelles choisies ou sur les différents paquets choisis, puis d'utiliser ces nouveaux coefficients en ligne pour les algorithmes de détection.

Le nombre de points initiaux nécessaires à cette modélisation est cependant supposé déjà disponible. Pour la décomposition, ce nombre est égal à la longueur du filtre passe bas permettant le calcul des coefficients d'approximation, et égal à la longueur de filtre passe haut permettant le calcul des coefficients de détail dans le cas de DWT. Il est, enfin, égal à la longueur du filtre utilisé pour le calcul des coefficients de chaque paquet dans le cas de WPT.

La phase de détection en temps réel est facilement réalisée dans le cas d'analyse directe sur les coefficients, car on ne calcule dans ce cas que les coefficients obtenus après décimation.

La phase de reconstruction complique beaucoup le problème de détection surtout en ce qui concerne l'insertion des zéros et les filtrages successifs du niveau d'échelle ou du paquet en cours jusqu'à l'arrivée au niveau du signal original.

Les coefficients d'ondelettes portent dans les échelles les plus élevées les informations utiles concernant les événements rapides qui correspondent aux hautes fréquences..Ils portent

également dans les échelles de niveaux les plus bas des informations relatives aux événements lents correspondant aux basses fréquences. Ils caractérisent de manière précise les événements. Les coefficients de chaque paquet dans le cas de décomposition en paquets d'ondelettes représentent l'information relative à tout le signal dans une bande de fréquence donnée. Il est tout à fait légitime de les utiliser directement pour le problème de détection dans l'EMG utérin.

4.3 Théorie du décalage des coefficients d'ondelettes

4.3.1 Support temporel des coefficients d'ondelettes

La représentation bidimensionnelle classique d'une décomposition en ondelettes est présentée sur la figure 4.1. La "colonne" constituée des pavés situés à la verticale d'un pavé de l'échelle d'approximation a_5 (un pavé sur l'échelle des détails d_5 , deux sur d_4 , jusqu'à 16 sur d_1) ne représente pas un intervalle temporel unique.

Sur chaque échelle de détails j de la décomposition, il y a $M / 2^j$ coefficients et le même nombre de pavés temps-fréquence pour un signal de longueur M . Il y a donc $M / 2^j$ ondelettes générées par ces coefficients. Cependant, la taille du support d'une ondelette est supérieure à 2^j points puisqu'elle est donnée par son ordre et donc par la longueur de la réponse impulsionnelle de son filtre associé.

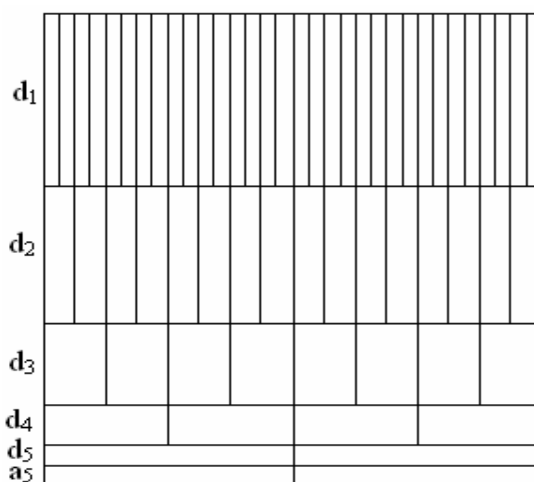


Figure 4.1. Pavage temps fréquence correspondant à chacun des niveaux de décomposition.

Les coefficients situés dans la même colonne du plan temps-échelle, mais à des échelles différentes, correspondent à des intervalles temporels ayant une longueur qui varie avec l'échelle. Il est nécessaire de calculer combien d'ondelettes successives à l'échelle j occupent le même intervalle temporel que celui occupé par l'ondelette de l'échelle $j + 1$, sachant que le décalage entre elles est de 2^j points. Considérons alors une ondelette à l'échelle $j + 1$. La taille de son support temporel est $T_{j+1} = (L - 1).(2^{j+1} - 1) + 1$ [Ran03], où L est la longueur de la réponse impulsionnelle du filtre associé (10 points pour l'ondelette Symlet no. 5). Considérons également une succession d'ondelettes à l'échelle j (de taille $T_j = (L - 1).(2^j - 1) + 1$) ayant le même instant de fin (ou de début) que l'ondelette située à l'échelle de plus basse fréquence $j + 1$. Comme $T_j < T_{j+1}$, il est nécessaire de calculer combien d'ondelettes successives à l'échelle j occupent le même intervalle temporel que celui occupé par l'ondelette de l'échelle $j + 1$, sachant que le décalage entre elles est de 2^j points. En notant n ce nombre, l'égalité d'intervalles temporels s'écrit [Ran03]:

$$T_{j+1} = T_j + n.2^j. \quad (4.1)$$

En remplaçant les expressions des tailles T_{j+1} et T_j , on obtient :

$$n = \frac{T_{j+1} - T_j}{2^j} = L - 1. \quad (4.2)$$

On peut donc avoir jusqu'à L ondelettes successives à l'échelle j responsables du même intervalle temporel qu'une seule ondelette de l'échelle $j + 1$. Autrement dit, les deux coefficients (pavés) de l'échelle j situés "au dessus" du coefficient de l'échelle $j + 1$ ne sont pas suffisants et il faut prendre encore $L - 2$ coefficients [Che05a], [Ran03].

Ce qui précède a son importance dans la recherche des coefficients responsables d'un intervalle temporel d'une transitoire dans un signal. Cette théorie s'applique de la même façon aux représentations des coefficients d'approximations aux différents niveaux d'échelles ou sur les différents paquets d'ondelettes.

Nous présentons au paragraphe suivant une méthode de calcul permettant la recherche du support temporel vrai d'un changement détecté à un niveau d'échelle donné ou pour un paquet d'ondelettes donné, afin d'en déduire ensuite l'estimation de la position de l'instant de changement sur le signal de départ.

4.3.2 Cas de la décomposition discrète en ondelettes

Un signal de départ peut être décomposé en approximations et en détails. Sur chaque niveau on obtient des coefficients filtrés et sous échantillonnés. Comme cela a été vu au paragraphe précédent, le calcul d'un coefficient à une échelle donnée nécessite L coefficients de l'échelle qui le précède. Par conséquent, chaque coefficient est influencé par tout un ensemble de points du signal original. Chaque coefficient calculé sur chaque niveau d'échelle est également décalé, et ce décalage est dû aux facteurs suivants:

- La décimation en passant d'un niveau à un autre.
- Le nombre de moments nuls des ondelettes utilisées (la longueur des réponses impulsionnelles des filtres passe bas et passe haut associés aux fonctions d'échelle et d'ondelettes).
- La forme de l'ondelette reliée à sa longueur et à sa symétrie.

Chaque coefficient peut être vu comme étant le résultat de l'estimation barycentrique de L coefficients de l'échelle précédente. Pour cette raison, nous avons introduit l'idée qu'une partie du décalage des coefficients est reliée à la position de barycentre "énergétique" du filtre en cours. Ce barycentre est obtenu pour chaque filtre par la formule suivante :

$$t_b^n = \frac{\sum_{m=1}^L m \cdot y_m^2}{\sum_{m=1}^L y_m^2} \quad (4.3)$$

où m est l'index de la position des coefficients du filtre, y_m est la valeur du coefficient à la position m , L est la longueur du filtre, et $t_b^n = t_b^g$ est la position du barycentre de $g(x)$ si n est impair. Dans l'autre cas $t_b^n = t_b^h$ est la position du barycentre de $h(x)$.

En examinant la figure 4.2, on voit que les positions des barycentres énergétiques des réponses impulsionnelles des filtres passe bas et passe haut sont 3.1495 et 1.8505, respectivement. Ces deux filtres sont associés à l'ondelette Symlet 2. Donc, si on veut connaître l'influence de l'instant 30 du signal original sur le premier niveau de décomposition, on fait la convolution du filtre passe haut avec la séquence numérique aux

instants 29, 30, 31 et 32. C'est donc le point d'indice 16 qui porte l'influence de l'instant 30 du signal original. On a, dans ce cas, un décalage de deux points, car la valeur la plus énergétique de la réponse impulsionnelle du filtre passe haut est la troisième en partant de la droite comme le montre la figure 4.2. De même le coefficient de détail d'indice 2 du niveau 4 porte l'influence du point à l'instant 9 du signal original comme le montre la figure 4.3, et on obtient un décalage de 23 points.

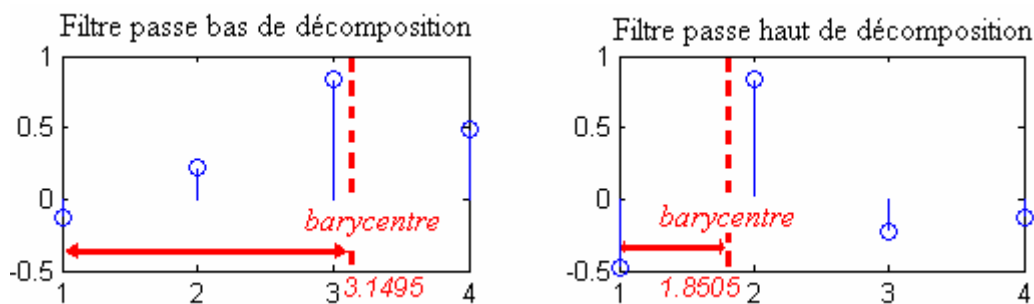


Figure 4.2. Réponses impulsionnelles des filtres passe bas et passe haut de l'ondelette Symlet no.2. Les positions des barycentres sont 3.1495 et 1.8505 respectivement.

La figure 4.3 présente l'effet du décalage aux différents niveaux d'échelle. L'ondelette choisie est l'ondelette Symlet 2. Les réponses impulsionnelles de ses deux filtres associés sont représentés sur la figure 4.2.

La problématique qui se pose alors est de calculer ce décalage, si notre but est de trouver le point (ou l'intervalle) du signal original représentant réellement un coefficient de détail ou d'approximation. Le calcul des coefficients de détail est différent de celui d'approximation car les filtres utilisés ne sont pas les mêmes dans les deux cas. Les corrections à appliquer dépendent donc du type de filtre.

Soit un coefficient d'approximation $a_j(k)$ à l'échelle j , et à la position k .

Pour ce type de coefficients, la procédure de correction est la suivante :

- 1- multiplier la position k par 2^j pour adapter l'échelle au signal original.
- 2- pour chaque translation du niveau j au niveau $j - 1$, calculer la distance entre les extrémités droites des filtres et la position du barycentre énergétique du filtre passe bas h .
- 3- multiplier chaque distance calculée à l'étape 2, par la puissance de 2 correspondant au niveau courant, afin de compenser le problème de décimation.

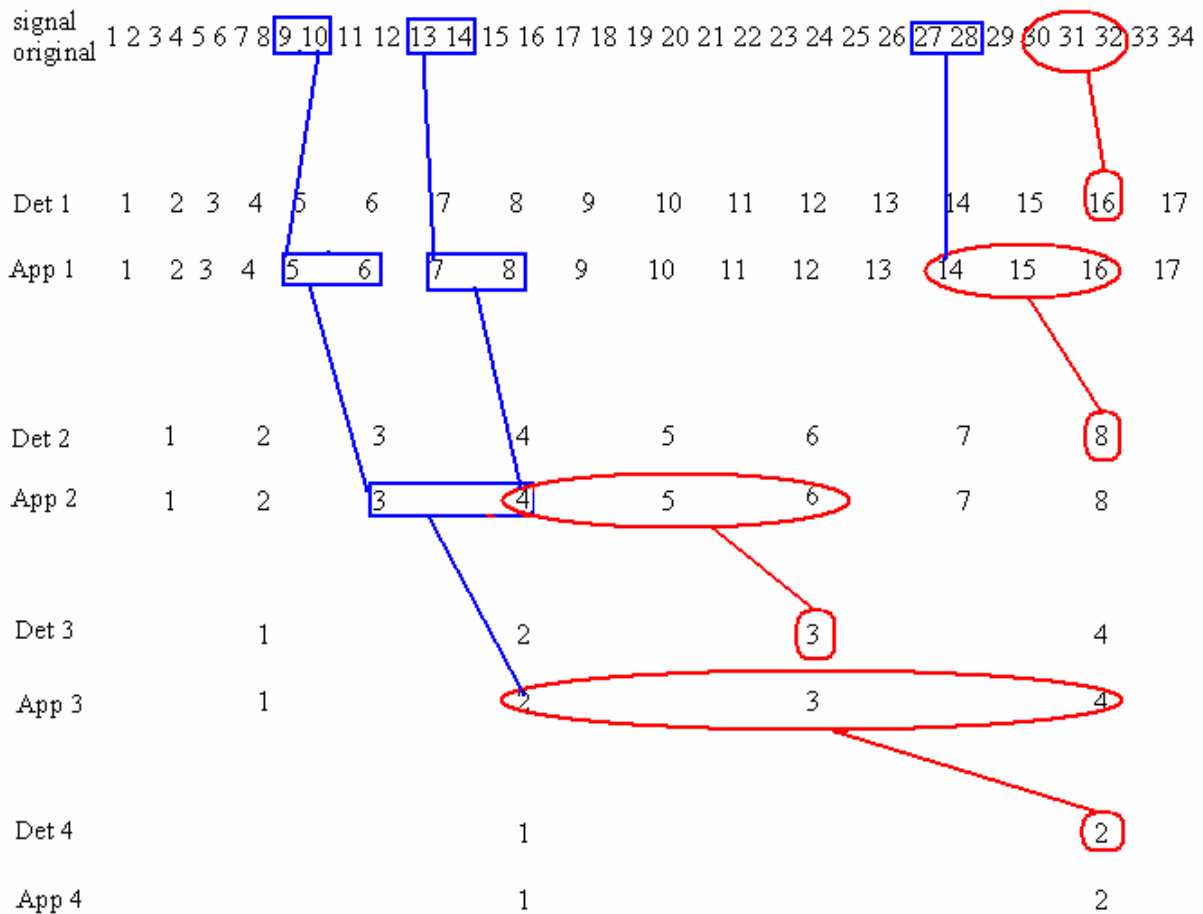


Figure 4.3. Illustration du décalage produit par le calcul des coefficients de détails et d'approximations. L'ondelette Symlet no.2 est utilisée.

La position correcte du coefficient $a_j(k)$ par rapport au signal original se traduit alors par la formule suivante:

$$t_j(k) = 2^j \cdot k - \sum_{i=0}^{j-1} 2^i \cdot (L - t_b^h) \quad (4.4)$$

ou encore :

$$t_j(k) = 2^j \cdot k - (2^j - 1)(L - t_b^h) \quad (4.5)$$

avec $(L - t_b^h)$ la longueur du filtre passe bas de l'extrémité droite au barycentre énergétique.

Soit un coefficient de détail $d_j(k)$ à l'échelle j , et à la position k .

Pour ce type de coefficient, la procédure est la suivante :

1- multiplier la position k par 2^j pour adapter l'échelle au signal original.

2- pour chaque translation du niveau j au niveau $j - 1$, calculer la distance entre les extrémités droites des filtres et la position du barycentre énergétique du filtre correspondant (h ou g).

3- multiplier chaque distance calculée à l'étape 2, par la puissance de 2 correspondant au niveau courant, afin de compenser le problème de décimation.

Alors la position correcte du coefficient $d_j(k)$ par rapport au signal original se traduit par la formule suivante:

$$t_j(k) = a_j(k)2^j - \left(\sum_{i=j-2}^0 ((L - t_b^h)2^i) + (L - t_b^g)2^{j-1} \right) \quad (4.6)$$

ou encore :

$$t_j(k) = d_j(k).2^j - (L - t_b^h)(2^{j-1} - 1) - (L - t_b^g).2^{j-1}. \quad (4.7)$$

avec $(L - t_b^h)$ et $(L - t_b^g)$ les longueurs des filtres passe bas et passe haut de leur extrémité droite à leur barycentre énergétique.

4.3.3 Cas de la décomposition en paquets d'ondelettes

La décomposition en paquets d'ondelettes est intéressante dans la mesure où seulement les paquets représentatifs du signal initial sont pris en considération. Le principe de décalage est le même que dans le cas de décomposition discrète en ondelettes, sauf que dans ce cas il faut prendre en considération tous les paquets amenant au paquet dont on veut calculer le décalage et par suite identifier les filtres nécessaires pour les faire entrer dans la définition de la correction du décalage.

Chaque nœud de l'arbre des paquets d'ondelettes correspond à une série de coefficients filtrés et sous échantillonnés. Chaque coefficient est le résultat d'une convolution avec le niveau précédent. Le calcul d'un coefficient d'un paquet d'ondelettes de niveau j et d'index $2n$ ou $2n+1$ nécessite l'emploi de L coefficients du paquet précédent de niveau $j-1$ et d'index n , L étant la longueur des filtres h et g .

Le décalage dans la position temporelle de chaque coefficient d'un paquet d'ondelettes est dû aux mêmes facteurs que ceux précisés dans le cas de décalage des coefficients d'ondelettes.

Chaque coefficient peut être vu comme étant le résultat de l'estimation barycentrique de L coefficients de l'échelle précédente. Rappelons que le barycentre énergétique pour chaque filtre est obtenu par la formule suivante :

$$t_b^n = \frac{\sum_{m=1}^L m \cdot y_m^2}{\sum_{m=1}^L y_m^2} \quad (4.8)$$

où m est l'index de la position des coefficients du filtre, y_m est la valeur du coefficient à la position m, L est la longueur du filtre, et $t_b^n = t_b^g$ est la position du barycentre de g(x) si n est impair. Dans le cas contraire $t_b^n = t_b^h$ est la position du barycentre de h(x).

La figure 4.4 illustre cette notion dans le cas d'une ondelette Symlet no. 3.

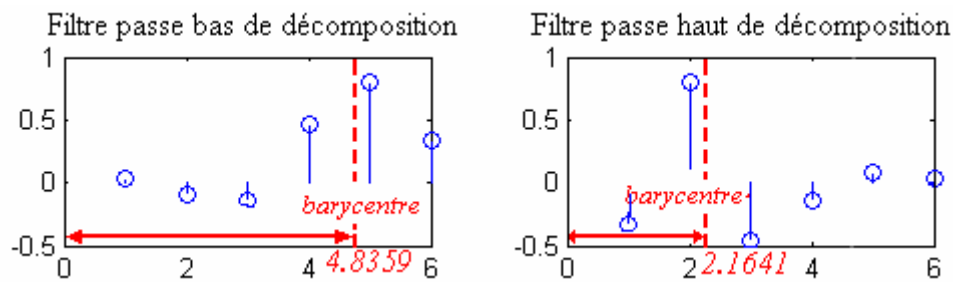


Figure 4.4. Réponses impulsionnelles des filtres passe bas et passe haut de l'ondelette Symlet no.3. Les positions des barycentres énergétiques sont 4.8359 et 2.1641 respectivement.

Donc si on veut connaître l'influence du point à l'instant 18 du signal original sur le paquet $W_{1,1}$ (figure 4.5), il faudrait considérer la convolution entre le filtre passe haut et la séquence numérique aux instants 17, 18, 19, 20, 21 et 22. C'est alors le point à l'instant 11 dans le paquet $W_{1,1}$ qui porte l'influence du point de l'instant 18 du signal original. Dans ce cas on a un décalage d'à peu près 4 points.

De même l'effet du coefficient à l'instant 11 du paquet $W_{1,1}$ se retrouve sur le coefficient no. 6 du paquet $W_{2,2}$ avec un décalage de 1 point, celui numéro 6 du paquet $W_{2,2}$ se trouve à l'instant 5 du paquet $W_{3,5}$ avec un décalage de 4 points, celui du numéro 5 du paquet $W_{3,5}$ à l'instant 3 du paquet $W_{4,10}$ avec un décalage d'un point. Le coefficient à la position 3 du paquet $W_{4,10}$ a un décalage de 30.82 points par rapport au coefficient no. 18 du signal original.

La correction appliquée à la position des coefficients des paquets d'ondelettes est alors la suivante :

Soit un coefficient $C_{j,n}(k)$ à l'échelle j , au paquet d'index n et à la position k .

La position correcte $t_{j,n}(k)$ du coefficient $C_{j,n}(k)$ par rapport à l'échelle précédente $j - 1$ est :

$$t_{j,n}(k) = k \cdot 2 - (L - t_b^n) \tag{4.9}$$

alors que la position correcte du coefficient $C_{j,n}(k)$ par rapport au signal original est :

$$t_{j,n}(k) = k \cdot 2^j - \sum_{i=0}^{j-1} \left(L - t_b^{\text{int}\left(\frac{n}{2^{j-i-1}}\right)} \right) 2^i. \tag{4.10}$$

avec $t_b^{\text{int}\left(\frac{n}{2^{j-i-1}}\right)} = t_b^g$ est la position du barycentre de g si $\text{int}\left(\frac{n}{2^{j-i-1}}\right)$ est impair et $t_b^{\text{int}\left(\frac{n}{2^{j-i-1}}\right)} = t_b^h$ la position du barycentre de h si $\text{int}\left(\frac{n}{2^{j-i-1}}\right)$ est pair.

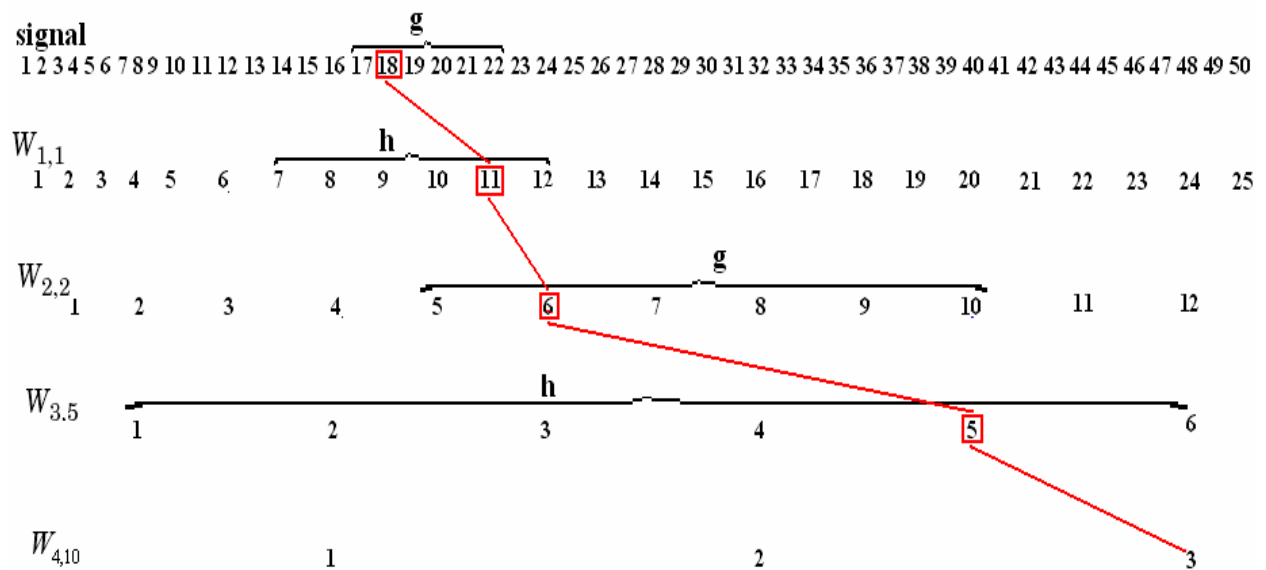


Figure 4.5. Illustration du décalage produit par la technique de calcul des coefficients des paquets d'ondelettes. L'effet du point no. 18 du signal original apparaît sur le coefficient no. 11 du paquet $W_{1,1}$, sur le coefficient no. 6 du paquet $W_{2,2}$, sur le coefficient no. 5 du paquet $W_{3,5}$ et sur le coefficient no. 3 du paquet $W_{4,10}$. L'ondelette Symlet no.3 est utilisée.

4.3.4 Synthèse

La méthode proposée pour la correction de décalage des coefficients a été expliquée. Elle a été basée principalement sur le calcul de la position du coefficient sur lequel est translatée l'influence d'un coefficient du signal original. Cette influence est produite par la convolution de la séquence par les filtres passe bas ou passe haut, et par la décimation. A chaque convolution la distance de la correction est calculée, puis multipliée par 2^j pour compenser la décimation. Les équations de correction dans le cas de décomposition discrète en ondelettes et dans le cas de décomposition en paquets d'ondelettes sont tout à fait similaires dans leur formulation.

Les deux ondelettes Symlet 2 et Symlet 3 ont été utilisées à titre d'illustration dans le paragraphe précédent pour expliquer le principe de décalage. Le choix définitif de l'ondelette reste à faire. Un des critères de choix de l'ondelette sera lié à l'erreur de localisation de l'instant de changement.

4.4 Algorithmes de détection de changements et application aux coefficients d'ondelettes

Dans le cadre de ce travail, l'objectif principal était de voir s'il était possible de détecter d'une façon effective la survenue d'événements dans l'EMG utérin en utilisant uniquement quelques paquets d'une décomposition en paquets d'ondelettes. Dans le chapitre précédent, nous avons établi une méthodologie de sélection de la meilleure base. Le présent chapitre a donc pour objectif de justifier ce terme de meilleure base. Des travaux antérieurs [Kha99], [Kha99a], [Kha97] se sont déjà intéressés de façon détaillée à la mise en place d'algorithmes de détection bien adaptés à l'EMG utérin. Dans ce chapitre, nous ne ferons que reprendre succinctement les différentes méthodes de détection et décrire d'une façon rapide l'algorithme retenu. D'une façon générale, la détection est abordée comme un test d'hypothèses résumé comme :

H_0 : pas de changement présent dans la séquence analysée

H_1 : il existe un changement dans la séquence analysée

Le changement est alors défini comme une modification des propriétés statistiques du signal.

Quand les paramètres des distributions associées sont connus, on peut définir un détecteur optimal basé sur le rapport de vraisemblance. Dans le cas où ces paramètres seraient inconnus, des auteurs ont utilisé le rapport de vraisemblance généralisé qui consiste à utiliser une estimation du maximum de vraisemblance à partir de l'estimation des paramètres. Dans le cas où les paramètres sont estimés au fur et à mesure de la progression de la détection, l'inconvénient majeur est qu'une nouvelle détection dépend des événements déjà détectés et identifiés. La performance de la méthode dépend donc en particulier de l'ordre d'apparition des événements.

Dans [Kha99a], une méthode alternative permet de réaliser la détection sans connaissance *a priori* de l'espace des hypothèses. Il sépare de plus les procédures de détection et de classification. Cette méthode est fondée sur une somme cumulée dynamique de rapports de vraisemblance généralisés entre deux distributions locales estimées avant et après l'instant courant t . Elle est appelée Somme Cumulée Dynamique (DCS). Cette méthode avait été appliquée soit sur le signal original après modélisation autorégressive, soit sur les signaux reconstruits après décomposition en ondelettes. Une amélioration a été faite dans [Fal02] conduisant à la Somme Cumulée Dynamique Modifiée (MDCS). Cette modification a pour but de détecter les changements lents. Dans cette nouvelle approche (MDCS), les signaux sont associés à une modélisation autorégressive. Nous allons décrire brièvement quelques types d'algorithmes de détection utilisables dans le cadre de notre travail.

4.4.1 Méthodes de détection de ruptures dans le cas d'un changement à paramètres connus

Les algorithmes de détection sont essentiellement fondés sur la théorie statistique des tests d'hypothèses. Parmi ces algorithmes, nous nous sommes intéressés aux algorithmes séquentiels qui nous ont semblé répondre au mieux à notre problème. Citons, par exemple, l'algorithme de somme cumulée (CUSUM) qui est l'un des algorithmes séquentiels de détection les plus efficaces dans le cas où les paramètres sont connus [Bas93].

4.4.1.1 Formulation du problème de détection

Le problème de détection consiste à détecter une éventuelle rupture dans les caractéristiques statistiques du signal observé \mathbf{x} et à estimer l'instant k de son apparition. Ce problème se

ramène à un test d'hypothèses. En effet, si nous disposons d'une suite d'observations x_1, x_2, \dots, x_n qui jusqu'à l'instant k suit une loi de probabilité f_{θ_0} (de paramètre θ_0) et ensuite est régie par une densité de probabilité f_{θ_1} (de paramètre θ_1), pour détecter un éventuel changement dans le paramètre θ , nous pourrions procéder, à chaque instant, au test d'hypothèse suivant:

$$H_0 : \theta = \theta_0, \text{ pour } 1 \leq i \leq n$$

contre

$$H_1 : \begin{cases} \theta = \theta_0, \text{ pour } 1 \leq i \leq k \\ \theta = \theta_1, \text{ pour } k+1 \leq i \leq n \end{cases}$$

(4.11)

Cette approche suppose évidemment connue la forme de la distribution f , ce qui sera le cas pour les signaux utérins.

Les méthodes de détection que nous allons utiliser sont fondées sur le maximum de vraisemblance, pour lequel a été montré qu'il constituait une statistique suffisante dans la théorie de la détection [Bor87], [Sap90], [Bas93]. Le logarithme du rapport de vraisemblance est défini par :

$$S(x_1, x_2, \dots, x_t) = \sum_{i=1}^t \ln \frac{f_{\theta_1}(x_i / x_{i-1}, \dots, x_1)}{f_{\theta_0}(x_i / x_{i-1}, \dots, x_1)} \quad (4.12)$$

Plusieurs algorithmes ont été développés pour résoudre le problème de détection dans ce cadre, parmi lesquels l'algorithme de Neyman-Pearson et l'algorithme de somme cumulée que nous allons présenter brièvement, car il sera à la base de la méthode retenue dans notre cas.

4.4.1.2 Algorithme de somme cumulée CUSUM

Formulation

La somme cumulée CUSUM est fondée sur le calcul récursif du logarithme du rapport de vraisemblance. La procédure CUSUM peut être considérée comme une séquence de tests répétés autour d'un point de changement k .

L'information à considérer se trouve alors dans la différence entre la valeur du rapport de vraisemblance et sa valeur minimale courante. La règle de décision est de comparer cette différence à un certain seuil, à chaque instant t , comme suit [Bas93]:

$$g_t = S_t - m_t \geq h \quad (4.13)$$

où

$$S_t = \sum_{i=1}^t s_i \quad (4.14)$$

$$s_i = \log \frac{f_{\theta_1}(y_i)}{f_{\theta_0}(y_i)} \quad (4.15)$$

$$m_t = \min_{1 \leq l \leq t} S_l \quad (4.16)$$

L'instant d'arrêt est :

$$t_a = \min\{t : g_t \geq h\} \quad (4.17)$$

L'instant de changement est défini alors par :

$$k = \max\{t < t_a : g_t = 0\} \quad (4.18)$$

La somme S_t fournit un test séquentiel de logarithme du rapport de vraisemblance. La propriété statistique essentielle de la méthode CUSUM est que les espérances du rapport de vraisemblance changent de signe après l'instant de rupture t_r :

$$E_{\theta_0}(s_i) < 0 \quad \text{pour } t < t_r$$

$$E_{\theta_1}(s_i) > 0 \quad \text{pour } t > t_r$$

Donc une variation du paramètre θ de θ_0 à θ_1 (et inversement) est reflétée par un changement de signe de la moyenne du logarithme du rapport de vraisemblance.

CUSUM en cas de changement de variance

Dans ce cas, on considère que les échantillons successifs suivent une loi gaussienne de moyenne nulle, et présentent des changements en variance uniquement. Ils sont supposés indépendants. Dans ce cas $\theta_0 = \sigma_0$ et $\theta_1 = \sigma_1$. Les densités de probabilité s'écrivent:

$$f_{\theta_0}(x_i) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{x_i^2}{2\sigma_0^2}} \quad (4.19)$$

$$f_{\theta_1}(x_i) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{x_i^2}{2\sigma_1^2}} \quad (4.20)$$

L'expression du logarithme de vraisemblance est alors:

$$s_i = \frac{1}{2} \left[\ln \frac{\sigma_1^2}{\sigma_0^2} + x_i^2 \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \right) \right] \quad (4.21)$$

4.4.2 Méthodes de détection des ruptures dans le cas de paramètres inconnus

En pratique, les paramètres des distributions ne sont généralement pas connus et beaucoup d'algorithmes ont été développés pour résoudre ce type de problème [Bas93].

4.4.2.1 Méthode du rapport de vraisemblance généralisé

Cette méthode est développée dans le cas où les échantillons sont indépendants. La somme des rapports de vraisemblance entre deux instants t_l et t est donnée par:

$$S_{t_l}^t = \sum_{i=t_l}^t s(x_i) = \sum_{i=t_l}^t \ln \frac{f_{\theta_1}(x_i)}{f_{\theta_0}(x_i)} \quad (4.22)$$

Lorsqu'on connaît θ_0 et θ_1 , l'instant de changement est estimé par k au sens du maximum de vraisemblance :

$$k = \arg \max_{t_1 \leq i \leq t} S_i^t \quad (4.23)$$

Si les paramètres θ_1 et θ_0 sont inconnus, ils sont remplacés par leur estimation au sens du maximum de vraisemblance :

$$\theta_1 = \arg \max_{\theta_1} S_{t_1}^t \quad \text{ou} \quad \theta_0 = \arg \max_{\theta_0} S_{t_1}^t \quad (4.24)$$

ce qui conduit finalement à la fonction de décision suivante :

$$g_t = \max_{1 \leq i \leq t} \sup_{\theta_1} S_i^t \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} h \quad (4.25)$$

Lorsque g_t dépasse un seuil h fixé *a priori*, H_1 doit être décidée. H_0 est décidée dans le cas contraire.

La double maximisation apparaissant dans la fonction de décision précédente est compliquée et coûteuse en temps de calcul. Elle est contradictoire avec la volonté de développement de méthodes capables de détecter séquentiellement d'éventuels changements. Plusieurs méthodes [Bra83] ont été proposées pour simplifier la méthode précédente.

4.4.2.2 Somme Cumulée Dynamique (DCS)

Principe

Soit $X = (x_1, \dots, x_t)$ le vecteur d'observation à l'instant t . Supposons que la distribution du processus X dépende d'un paramètre θ . Un changement peut affecter la variance, la distribution fréquentielle du signal ou les deux en même temps. La méthode de somme cumulée dynamique est une méthode construite pour la détection des changements dans des signaux dans le cas où les paramètres sont inconnus, et lorsqu'on veut suivre des changements locaux dans les signaux.

Cette approche de détection, appelée la Somme Cumulée Dynamique (DCS : Dynamic Cumulative Sum), peut être considérée, comme la méthode CUSUM, comme une séquence

répétée autour du point de changement k . Elle est fondée sur la somme cumulée locale des rapports de vraisemblance entre deux distributions locales estimées autour de l'instant courant t . Ces deux distributions locales D_a^t ("after t ") et D_b^t ("before t ") sont estimées en utilisant deux fenêtres de longueur N (figure 4.6) avant et après l'instant t comme suit :

$x_i; i = \{t - N, \dots, t - 1\}$ suit une loi de probabilité de densité $f_{\theta_0}(x_i)$

$x_i; i = \{t + 1, \dots, t + N\}$ suit une loi de probabilité de densité $f_{\theta_1}(x_i)$

Les paramètres de la distribution $D_b^t, \hat{\theta}_b$ sont estimés à partir de N points avant l'instant courant t tandis que ceux de la distribution $D_a^t, \hat{\theta}_a$ sont estimés à partir de N points après l'instant t .

A l'instant t , nous définissons DCS comme la somme des logarithmes des rapports de vraisemblance à partir du début du signal jusqu'à l'instant t [Kha99b] :

$$DCS(D_a^t, D_b^t) = \sum_{i=1}^t \ln \frac{f_{\hat{\theta}_a}^i(x_i)}{f_{\hat{\theta}_b}^i(x_i)} = \sum_{i=1}^t \hat{s}_i \quad (4.26)$$

avec $\hat{s}_i = \ln \frac{f_{\hat{\theta}_a}^i(x_i)}{f_{\hat{\theta}_b}^i(x_i)}$. Le logarithme du rapport de vraisemblance a un caractère local dans la

mesure où les paramètres des deux distributions sont ré-estimés à chaque pas dans les deux fenêtres de N points autour du point courant t .

Il a été démontré dans [Kha99b] que la fonction DCS arrive à son maximum au temps de changement k . Un exemple de l'évolution de DCS est présenté sur la figure 4.6.

La fonction de détection utilisée pour estimer l'instant de changement est exprimée par:

$$g_t = \max_{1 \leq i \leq t} [DCS(D_a^i, D_b^i)] - DCS(D_a^t, D_b^t) \quad (4.27)$$

L'instant d'arrêt est:

$$t_a = \inf \{t : g_t \geq h\} \quad (4.28)$$

h étant un seuil fixe.

Le vrai instant de changement est estimé par :

$$k = \sup \{t < t_a : g_t = 0\} \quad (4.29)$$

La formulation précédente est une formulation générale de DCS. C'est cet algorithme qui sera utilisé dans la suite pour l'évaluation des performances de l'algorithme de sélection de la meilleure base.

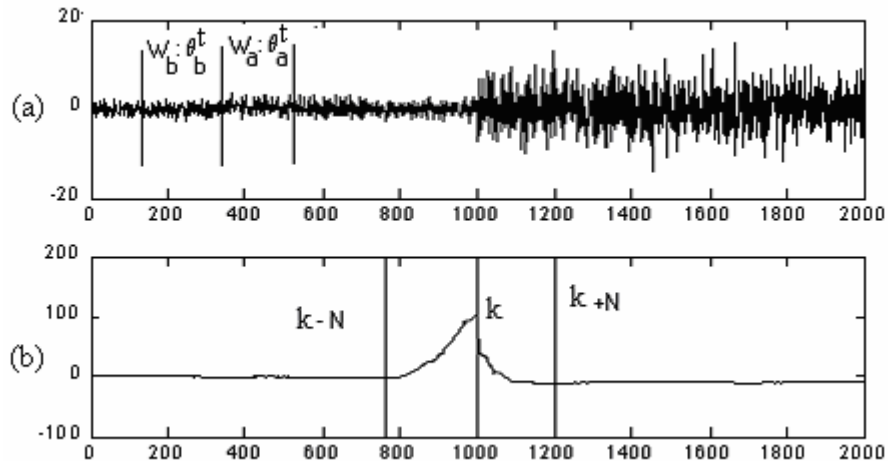


Figure.4.6 – (a) Exemple de signal contenant un seul point de changement k . (b) Evolution de la somme cumulée dynamique autour du point de changement. Axes des abscisses : nombre de points. Axes des ordonnées : unités arbitraires

4.4.3 Fusion des instants de changement

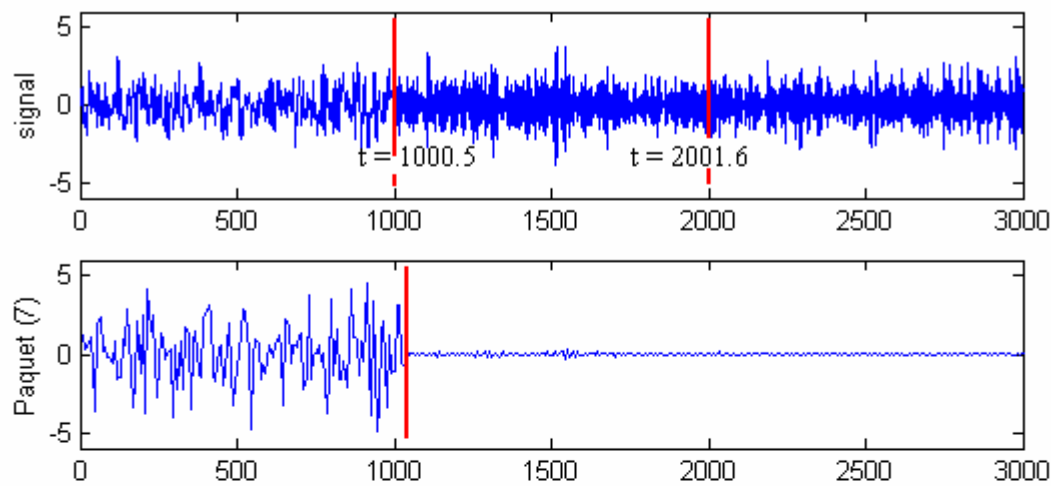
Puisque l'algorithme de détection choisi (quel qu'il soit) est appliqué sur chacun des paquets sélectionnés par l'algorithme du choix de la meilleure base, il y a nécessité de fusionner les différents instants de détection identifiés sur chaque paquet sélectionné pour obtenir un seul instant de changement pour une rupture donnée, (ou plus exactement un intervalle de changement).

L'algorithme de fusion proposé est le suivant (toutes les valeurs sont données en nombre de points) :

- Chaque instant de changement $t_{j,n}(k)$ détecté sur un paquet d'ondelette au niveau j et d'index n (de façon similaire pour les instants de changement sur les coefficients de détail ou d'approximation) est considéré comme un intervalle $[t_{j,n} - 0.5, t_{j,n} + 0.5]$. Ceci prend en compte la résolution temporelle du niveau du paquet d'ondelette.

- Chaque intervalle de détection ci-dessus est transformé vers sa position correcte par rapport au signal initial. Cette correction est fondée sur la méthode de correction des instants de changements décrite à la section 4.3.2 (décomposition en ondelettes) et 4.3.3 (décomposition en paquets d'ondelettes).
- Tous les intervalles présentant une zone temporelle commune (superposition partielle ou totale) au signal initial sont considérés liés au même changement.
- L'instant de changement correspondant est calculé comme étant la position du barycentre de tous les intervalles superposés.

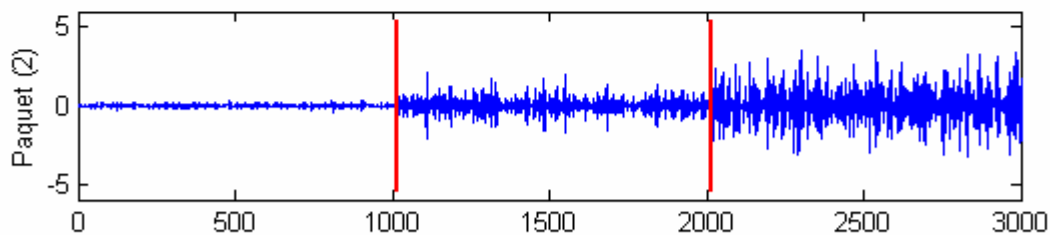
Cette procédure est illustrée à la figure 4.7.



$$t(\text{paquet}) = 129 \rightarrow t = 129 \cdot (2^3) = 1032$$

$$[t - 0.5(2^3) = 1028, t + 0.5(2^3) = 1036]$$

$$[1028 - 31.85 = 996.15, 1036 - 31.85 = 1004.2]$$



$$t(\text{paquet}) = 503 \rightarrow t = 503 \cdot (2) = 1006$$

$$[t - 0.5(2) = 1005, t + 0.5(2) = 1007]$$

$$[1028 - 4.45 = 996.15, 1036 - 4.45 = 1002.6]$$

$$t(\text{paquet}) = 1003 \rightarrow t = 1003 \cdot (2) = 2006$$

$$[t - 0.5(2) = 2005, t + 0.5(2) = 2007]$$

$$[2005 - 4.45 = 2000.6, 2007 - 4.45 = 2002.6]$$

Figure 4.7. Illustration de la correction et de la fusion des instants détectés sur les paquets sélectionnés.

Un signal synthétique de la base SIM2 est décomposé en paquets d'ondelettes. Rappelons que ce signal est formé de 3 segments de 1000 points chacun correspondant à différentes bandes fréquentielles dont les bornes sont définies par référence à la moitié de la fréquence d'échantillonnage $F_e/2$ ($[0.01 ; 0.25]$, $[0.26 ; 0.5]$, $[0.51 ; 0.75]$). La longueur totale du signal en nombre de points est 3000. Les paquets 7 et 2 sont choisis comme étant les paquets les plus significatifs en détection après application de l'algorithme du choix de la meilleure base décrit au chapitre 3. Les instants de changement sur chaque paquet sont estimés par DCS, puis les intervalles de détection qui y correspondent sont corrigés en utilisant l'équation 4.10 et finalement les intervalles de détection corrigés sont fusionnés pour obtenir l'instant du changement adapté au signal original. Dans ce cas, l'ondelette utilisée est Symlet 5. Notons que le choix de l'ondelette est arbitraire jusqu'à maintenant. Nous discuterons du meilleur choix de l'ondelette dans les paragraphes suivants. Comme les paquets utilisés sont 7 et 2 alors les distances de correction sont 31.85 et 4.45 respectivement, elles sont calculées en utilisant la deuxième partie de l'équation (4.10). Les vrais instants de changements sont 1000 et 2000 alors que ceux calculés par notre système sont 1000.5 et 2001.6 (en nombre de points).

4.5 Validation et performance des méthodes

4.5.1 Performance de la correction de décalage

Pour montrer l'efficacité de calcul des formules de décalage, les résultats que nous présentons sont fondés sur une étude en simulation. La simulation consiste à générer un millier des signaux, en créant des segments de bruit blanc gaussien, avec un changement de variance de 1 à 4. L'instant de changement est fixé à 2000 (numéro du point). La base des signaux de référence est la base SIM3 (annexe A).

DWT est appliquée sur les signaux simulés. La décomposition est appliquée jusqu'au niveau 5.

Puisque dans ce cas nous connaissons les paramètres des distributions, ce sont les algorithmes de détection du maximum de vraisemblance et CUSUM à paramètres connus dans le cas de changement de variance (voir section 4.4.1) qui sont appliqués ici sur chaque niveau de détail [Che05a]. Les instants de détection sont calculés puis corrigés selon la procédure expliquée à la section 4.3.2.

Les figures 4.8, 4.9 et 4.10 illustrent l'effet de décalage pour trois types d'ondelettes. Elles représentent les histogrammes des instants de rupture avant et après correction pour les coefficients de détail des niveaux d'échelle 2, 4 et 5.

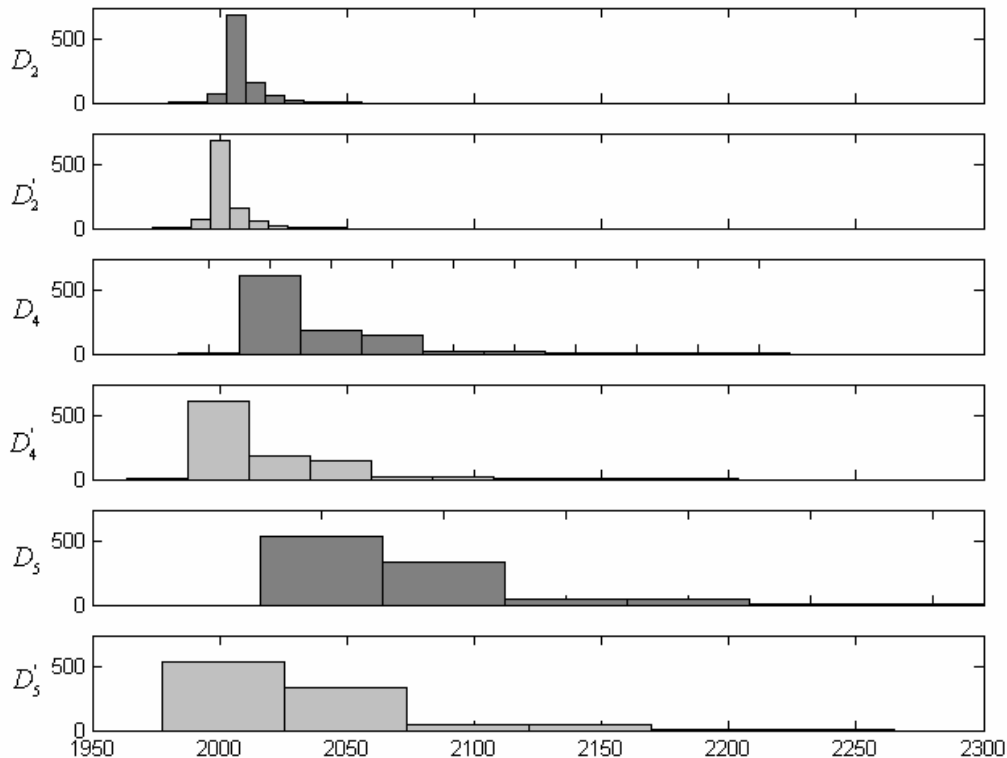


Figure 4.8. Histogrammes des instants de rupture sur les niveaux d'échelles 2, 4 et 5. L'ondelette utilisée ici est Daubechie no.3 (db3). Les histogrammes gris foncés correspondent aux instants avant correction alors que les histogrammes en gris clair correspondent aux instants après correction. Méthode de détection : maximum de vraisemblance.

Ces instants ont été obtenus en appliquant la méthode du maximum de vraisemblance sur chacun des 1000 signaux déjà décrits au début de ce paragraphe (nous avons vérifié que nous obtenions les mêmes résultats en appliquant l'algorithme de somme cumulée CUSUM). La figure 4.10 montre un autre exemple de recalage en présentant, comme les figures précédentes, les histogrammes des instants de rupture avant et après correction pour les coefficients de détail, mais cette fois pour l'ensemble des niveaux d'échelle de 1 à 5.

Les histogrammes en gris foncé correspondent toujours aux instants de rupture avant la correction et les histogrammes en gris clair aux instants de rupture obtenus après correction.

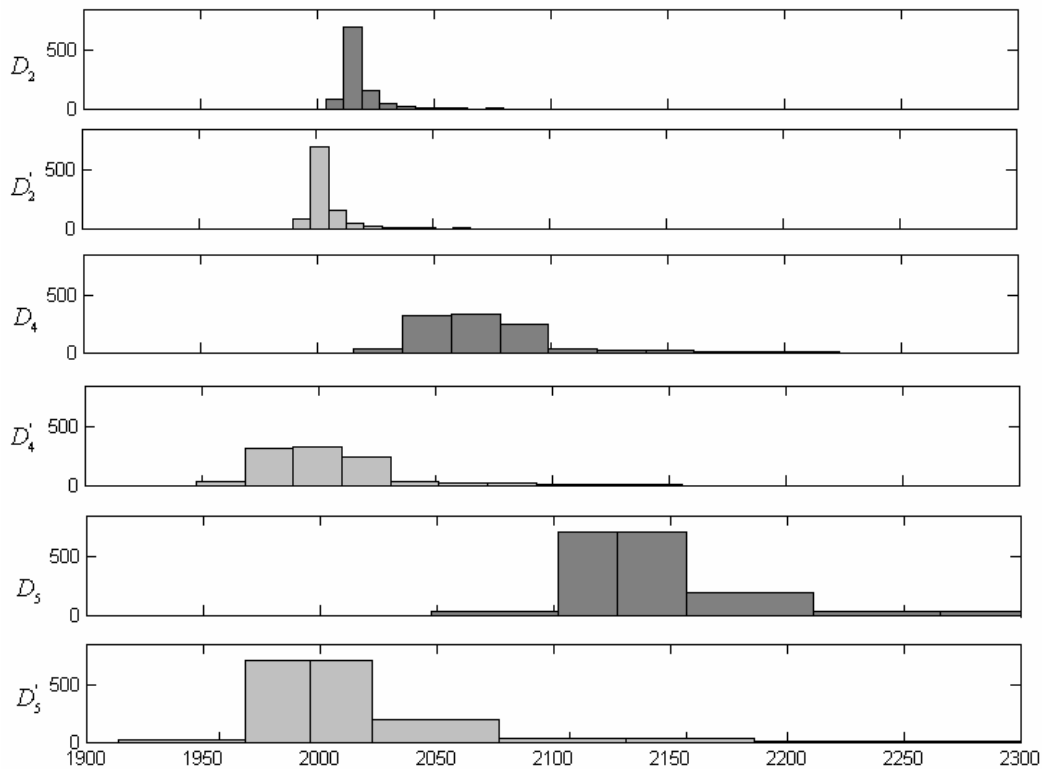


Figure 4.9. Histogrammes des instants de rupture sur les niveaux d'échelles 2, 4 et 5. L'ondelette utilisée est Symlet no.5. Les histogrammes gris foncés correspondent aux instants avant correction alors que les histogrammes en gris clair correspondent aux instants après correction. Méthode de détection : maximum de vraisemblance.

L'ondelette utilisée à la figure 4.8 est l'ondelette de Daubechie 3 (db3) qui a un support égal à $(2N-1 = 2.3-1 = 5)$ et une longueur de réponse impulsionnelle égale à 6. L'ondelette utilisée à la figure 4.9 est l'ondelette Symlet 5 (sym5) qui a un support égal à $(2.5-1 = 9)$ et une longueur de réponses impulsionnelles égale à 10. Enfin, l'ondelette utilisée à la figure 4.10 est l'ondelette de Coiflet 5 (coif5) dont le support égal à $(6N-1 = 6.5-1 = 29)$ et la longueur de sa réponse impulsionnelle est égale à 30. Les réponses impulsionnelles des ondelettes db3, sym5 et coif5 sont données à la figure 4.11.

Les différences selon l'ordre et le type de l'ondelette se voient clairement. Il a déjà été précisé que le décalage est influencé par l'ordre de l'ondelette utilisée en décomposition et par conséquent par la longueur de ses filtres associés. Plus l'ordre de l'ondelette et l'échelle de décomposition sont grands, plus le décalage sera important. De plus, la position du barycentre énergétique influence beaucoup le décalage aux différents niveaux. Toutes ces influences se manifestent clairement à partir des illustrations des figures 4.8, 4.9 et 4.10.

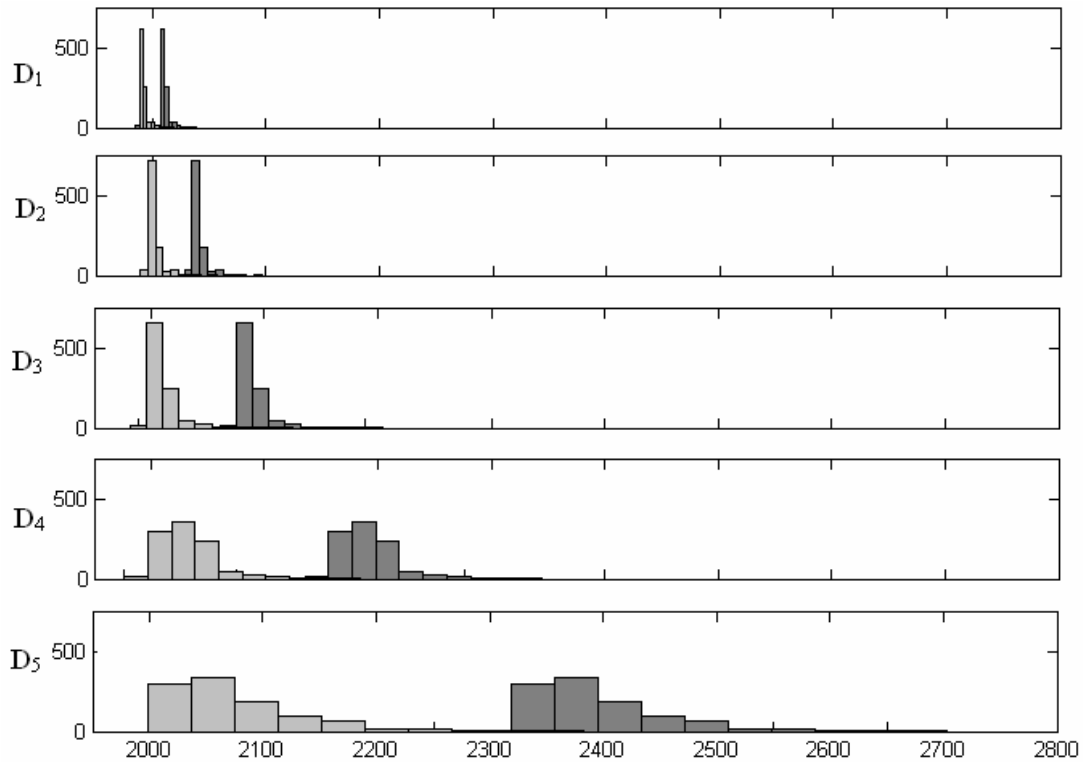


Figure 4.10. Histogrammes des instants de rupture sur les niveaux d'échelles 1, 2, 3, 4 et 5. L'ondelette utilisée est Coiflet no.5 (Coif5). Les histogrammes gris foncés correspondent aux instants avant correction alors que les histogrammes en gris clair correspondent aux instants après correction. Méthode de détection : maximum de vraisemblance.

Ce qui a été fait pour les niveaux d'échelles peut se faire de façon équivalente pour la décomposition en paquets d'ondelettes. Pour montrer l'efficacité du calcul des formules de décalage dans ce cas, les mêmes calculs que ceux du paragraphe précédent ont été répétés. Dans ce cas, le calcul s'est effectué sur tous les paquets d'ondelettes. La figure 4.12 montre l'historgramme du temps de changement avant et après la correction sur des paquets d'ondelettes arbitrairement choisis. C'est l'ondelette Symlet 5 qui a été choisie dans ce cas. Comme les distributions sont connues, la méthode de maximum de vraisemblance a été utilisée ici aussi pour calculer les instants de changement.

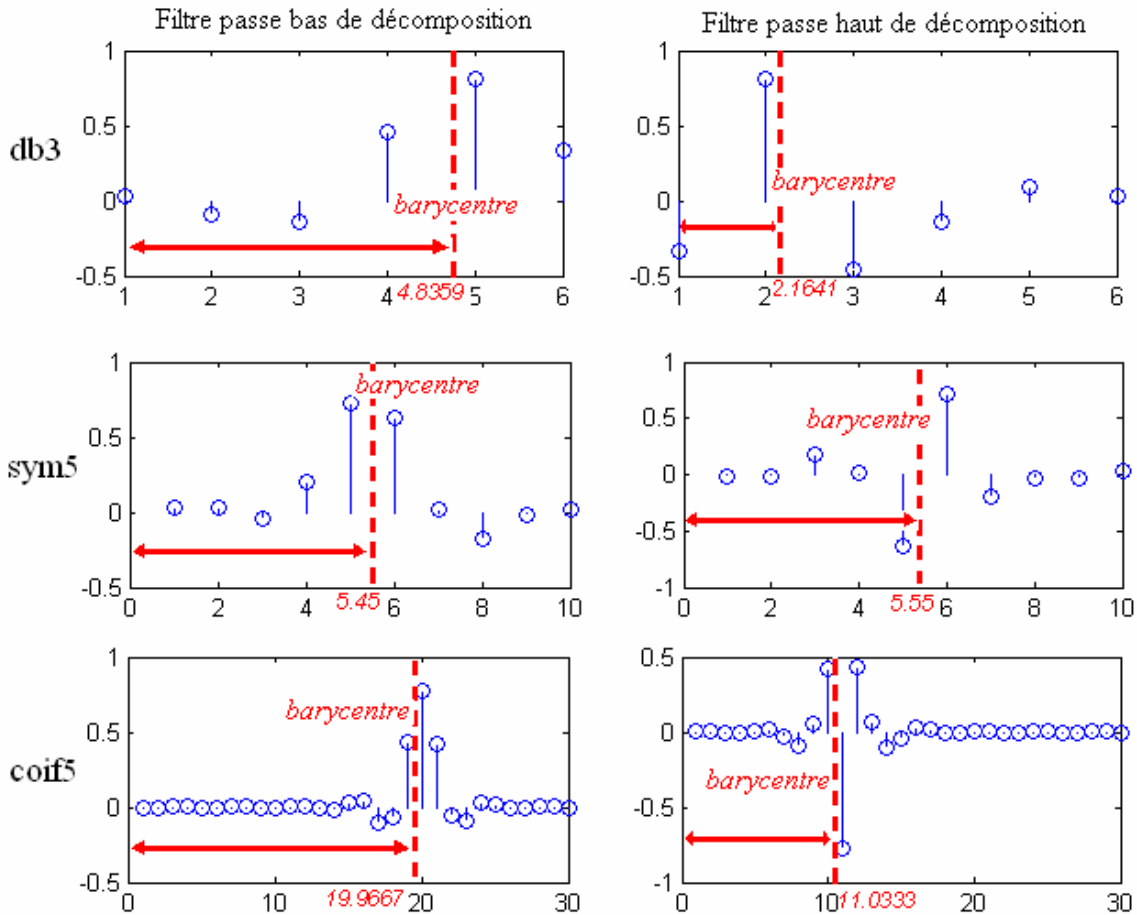


Figure 4.11. Réponses impulsionnelles des filtres passe bas et passe haut des ondelettes db3, sym5 et coif5, respectivement.

Du point de vue quantitatif, nous avons rassemblé dans les tableaux 4.1, 4.2 et 4.3 la moyenne des instants de changement avant et après correction, ainsi que les différences entre les instants de changements corrigés et la valeur de vrai instant de changement (2000), dans le cas de la décomposition en niveaux d'échelles. Pour les paquets d'ondelettes, le nombre important de tableaux nous a incités à les présenter en annexe D.

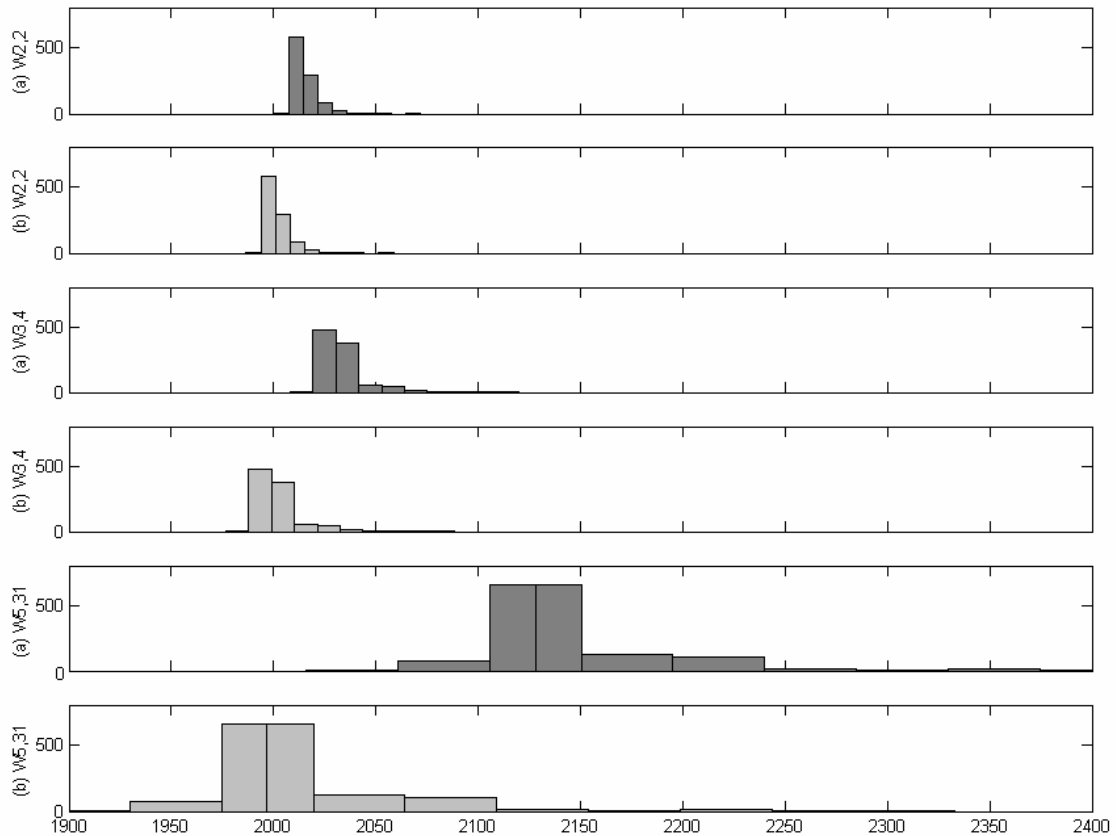


Figure 4.12. Histogrammes des instants de rupture sur les paquets $W_{2,2}$, $W_{3,4}$ et $W_{5,31}$ avant (a) et après la correction (b). L'ondelette utilisée est Symlet no.5. L'instant réel du changement est 2000. Méthode de détection : maximum de vraisemblance.

db3	t	$t.2^j$	t_c	$t_c - 2000$
D1	1001.1	2002.2	1998.4	-1.6099
D2	501.9	2007.8	2001.6	1.6199
D3	252.3	2018.7	2007.9	7.8913
D4	127.5	2039.9	2019.8	19.7542
D5	65	2079.7	2041	40.952

Tableau 4.1. Pour chaque niveau de détail (D1 à D5), colonne 1 : instant de changement détecté, colonne 2 : correction simple liée à la décimation, colonne 3 : correction de décalage, colonne 4 : écart à la vraie valeur. Résultats dans le cas de l'ondelette db3.

Sym5	t	$t.2^j$	t_c	$t_c - 2000$
D1	1003	2005.9	2001.5	1.478
D2	503.9	2015.8	2002.2	2.238
D3	254.2	2034	2002.2	2.2018
D4	129.3	2068.5	2000.4	0.3776
D5	67.1	2147.4	2006.4	6.409

Tableau 4.2. Résultats dans le cas de l'ondelette Sym5.

coif5	t	t.2^j	t_c	t_c -2000
D1	1005.6	2011.3	1992.3	-7.6807
D2	510	2040.1	2001.1	1.0747
D3	262.4	2098.9	2019.7	19.7453
D4	138.4	2214.4	2055	54.9505
D5	73.34	2347.8	2070.8	70.8

Tableau 4.3. Résultats dans le cas de l'ondelette Coif5.

Les tableaux 4.4 et 4.5 complètent les résultats précédents en les étendant à d'autres types d'ondelettes. Ils présentent les écarts types et les moyennes des différences entre les instants réels de changement et les instants obtenus après la correction dans les deux cas de décomposition discrète en ondelettes et de décomposition en paquets d'ondelettes (moyenne et écart type de la dernière colonne des tableaux 4.1, 4.2 et 4.3 dans le cas de décomposition en ondelettes, et de la dernière colonne des tableaux de l'annexe D dans le cas de décomposition en paquets d'ondelettes). Les résultats les plus satisfaisants sur la moyenne sont obtenus pour l'ondelette Symlet 5, surtout dans le cas de décomposition en paquets d'ondelettes. Les valeurs des écarts types sont en général très larges et augmentent avec le nombre de moments des ondelettes et par conséquent avec la longueur des filtres associés. Le meilleur résultat en terme d'écart type a été obtenu également pour l'ondelette Symlet no.5. De ce qui précède, nous avons décidé d'utiliser pour toute la suite du travail l'ondelette Symlet no.5.

Nom de l'ondelette	Ecart type	Moyenne
Symlet 2	9.7702	7.9775
Symlet 5	2.2901	2.5409
Symlet 8	5.2676	-2.5454
Coiflet 2	20.3584	15.4567
Coiflet 3	32.7656	23.7081
Coiflet 4	42.5372	32.1862
Daubechies 3	17.2727	13.7215
Daubechies 6	42.1459	19.9365

Tableau 4.4. Ecarts types et moyennes des différences entre l'instant réel de changement (2000) et les instants de détection après correction. Les résultats sont donnés pour une moyenne de 1000 signaux simulés décomposés sur les niveaux d'échelles 1, 2, 3, 4 et 5.

Nom de l'ondelette	Ecart type	Moyenne
Symlet 2	15.3963	6.3548
<i>Symlet 5</i>	<i>3.6672</i>	<i>4.7210</i>
Symlet 8	15.2459	5.3
Coiflet 2	33.8324	5.5871
Coiflet 3	56.1291	4.7452
Coiflet 4	78.3507	4.7508
Daubechies 3	30.9861	5.47
Daubechies 6	78.8124	4.045

Tableau 4.5. Ecarts types et moyennes des différences entre l'instant réel de changement (2000) et les instants de détection après correction. Les résultats sont donnés pour une moyenne de 1000 signaux simulés après décomposition en 62 paquets d'ondelettes jusqu'au niveau 5.

4.5.2 Performance des algorithmes de détection sur les meilleurs paquets choisis après correction et fusion des instants de changement

4.5.2.1 Résultats sur les signaux simulés

Les mêmes signaux simulés que ceux utilisés au chapitre précédent et appartenant à la base de signaux SIM1 sont utilisés. Chacun de ces 1000 signaux est composé de 4 segments à différentes bandes fréquentielles définies à partir de la moitié de la fréquence d'échantillonnage $F_e/2$ ([0.01 ; 0.4], [0.01 ; 0.6], [0.01 ; 0.8], et [0.01 ; 0.999] Hz). Chaque segment est composé de 2000 échantillons et chaque signal contient 3 instants de changements : 2000, 4000, et 6000. La longueur totale de chaque signal est de 8000 (nombre de points). Après application de l'algorithme de sélection de la meilleure base (paragraphe 3.6.3) sur l'ensemble d'apprentissage G1SIM1 de la base SIM1, les paquets 6, 11, et 12 ont été retenus comme étant la meilleure base incomplète pour cet ensemble d'apprentissage de signaux synthétiques. Les étapes du choix de la meilleure base sont illustrées sur la figure 4.13 (voir section 3.7.1).

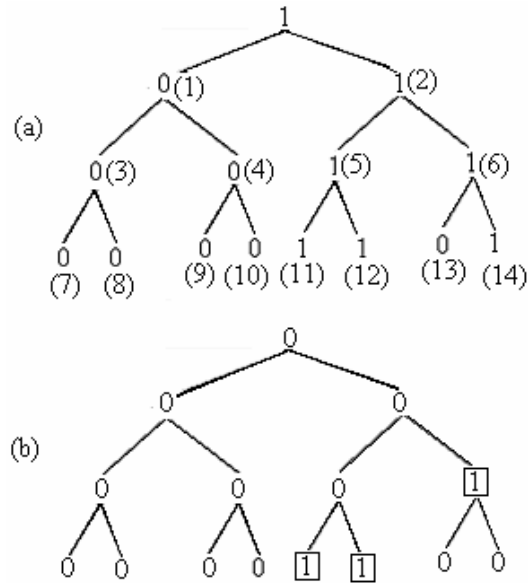


Figure 4.13. (a) Caractérisation de chacun des paquets selon les valeurs de D_{\max} (le numéro de chaque paquet est indiqué à l'intérieur des parenthèses). (b) Sélection de la meilleure base (incomplète dans ce cas).

Après la sélection de la meilleure base, l'algorithme de somme cumulée dynamique DCS est appliqué sur les coefficients des paquets choisis obtenus à partir de l'ensemble test G2SIM1 composé de 1000 signaux construits de la même façon que l'ensemble d'apprentissage G1SIM1. Pour montrer la performance de l'algorithme DCS en l'appliquant sur les coefficients, les probabilités de bonne détection et de fausse alarme ont été calculées. Les instants de changement utilisés pour calculer les probabilités de bonne détection et de fausse alarme ont été obtenus après correction et fusion des temps de changements évoqués aux paragraphes 4.3.3 et 4.4.3 respectivement. Les probabilités de bonne détection et de fausse alarme sont présentées dans le tableau 4.6.

Performance de l'algorithme de sélection de la meilleure base

On aborde dans cette section la performance de l'algorithme de sélection de la meilleure base après correction, fusion des instants de changement, et application de l'algorithme de détection.

Pour démontrer l'efficacité de la réduction du nombre de paquets après application de l'algorithme de choix de la meilleure base, le même algorithme de détection est appliqué à tous les paquets sélectionnés à la première étape de l'algorithme de sélection de la meilleure base (section 3.6.3). Ce sont les paquets à 1 de la figure 4.13.a (paquets 2, 5, 6, 11, 12, et 14).

Les 1000 signaux de l'ensemble test G2SIM1 ont été utilisés. Les nouvelles probabilités de bonne détection et de fausse alarme ont été calculées (tableau 4.6). Elles sont très proches de celles obtenues après réduction (paquets 6, 11 et 12). Ce résultat nous incite à confirmer que l'information utile a été conservée après la sélection de la meilleure base. La figure 4.14 illustre la détection des instants de changement pour un signal de la base SIM1.

	Avant réduction	Après réduction
Bonne détection	0.9652	0.953
Fausse alarme	0.0535	0.047

Tableau 4.6. Probabilités de bonne détection et de fausse alarme des signaux simulés avant et après application de l'algorithme du choix de la meilleure base pour la détection.

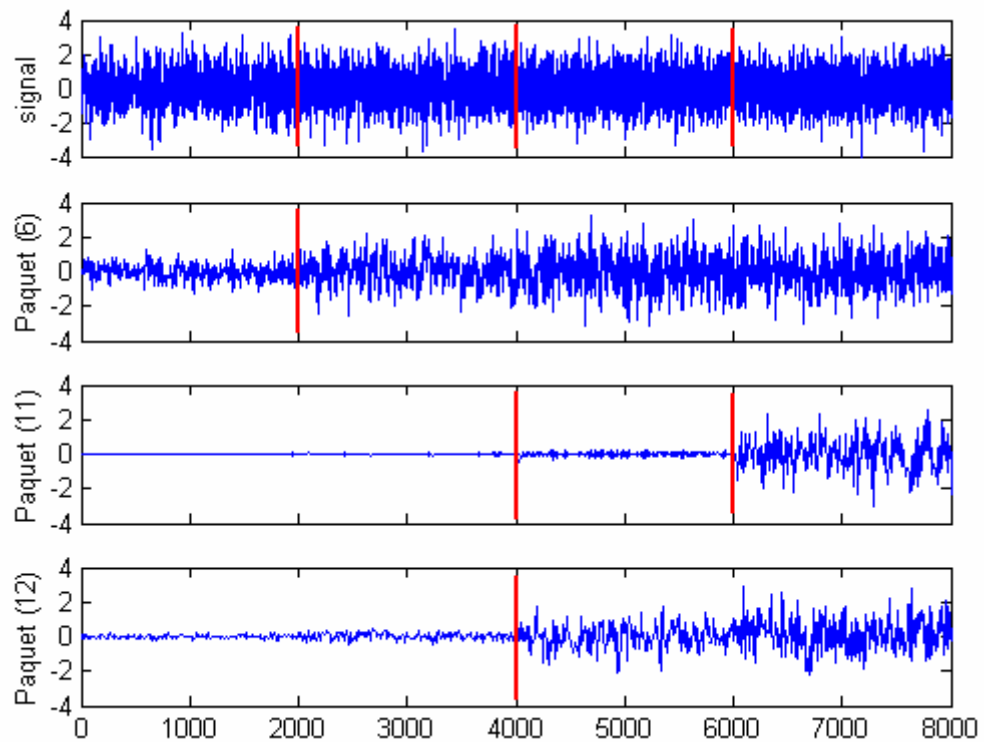


Figure 4.14. Signal simulé extrait de la base des signaux SIM1, décomposé sur les paquets 6, 11 et 12. Les tracés des coefficients de ces paquets sont illustrés.

4.5.2.2 Résultats sur les signaux réels

Pour les signaux EMG utérins ($F_e = 16$ Hz), l'ensemble test G2REELSIM de la base des signaux REELSIM (décrite à l'annexe A) est utilisé. Rappelons que cet ensemble est constitué

de 100 signaux. Chaque signal est composé de 10 événements de différents types (contractions, ondes d'Alvarez, ondes LDBF, mouvements actifs fœtaux, etc.) identifiés par un expert. Les événements ont été contaminés par du bruit blanc (SNR = 10 db) [Che05c]. Après application de la première étape de l'algorithme de sélection de la meilleure base (paragraphe 3.6.3) sur l'ensemble d'apprentissage G1REELSIM, les paquets à 1 de la figure 4.15.a (paquets 3, 4, 7 et 8) ont été sélectionnés. Après avoir effectué les étapes de réduction dans l'algorithme de sélection de la meilleure base, les paquets 7 et 8 (figure 4.15.b) ont été retenus comme étant la meilleure base incomplète pour cet ensemble d'apprentissage des signaux réels (voir section 3.7.2) [Che06].

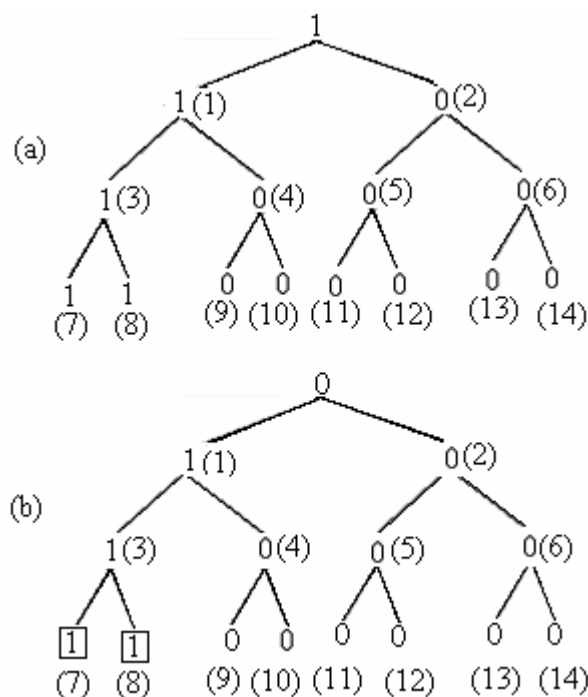


Figure 4.15. (a) Caractérisation de chacun des paquets selon les valeurs de D_{\max} (le numéro de chaque paquet est indiqué à l'intérieur des parenthèses). (b) Sélection de la meilleure base pour l'EMG utérin.

Après la sélection de la meilleure base, l'algorithme de somme cumulée dynamique DCS est appliqué sur les coefficients des paquets choisis. Ces paquets sont liés aux signaux de l'ensemble test G2REELSIM. Pour quantifier la performance de l'algorithme DCS en l'appliquant sur les coefficients, les probabilités de bonne détection et de fausse alarme ont été calculées. Les instants de changement utilisés pour calculer les probabilités de bonne détection et de fausse alarme ont été obtenus après correction et fusion des temps de changements comme évoqué aux sections 4.3.3 et 4.4.3 respectivement. Les probabilités de

bonne détection et de fausse alarme calculées après réduction sont présentées dans le tableau 4.7.

Comme dans le cas des signaux simulés, pour démontrer l'efficacité de la réduction du nombre de paquets après application de l'algorithme de choix de la meilleure base, l'algorithme de détection est appliqué à tous les paquets sélectionnés à la première étape de l'algorithme de sélection de la meilleure base. Ce sont les paquets 1, 3, 7 et 8. Les 100 signaux de l'ensemble test G2REELSIM ont été utilisés. Les nouvelles probabilités de bonne détection et de fausse alarme ont été calculées (tableau 4.7). Elles sont très proches de celles obtenues après réduction (paquets 7 et 8). Ceci veut dire que l'information utile a été conservée après la réduction de nombre des paquets [Che06]. La figure 4.16 illustre la détection sur un exemple choisi parmi les signaux utilisés de la base REELSIM.

	Avant réduction	Après réduction
Bonne détection	0.9989	0.9878
Fausse alarme	0.0652	0.0545

Tableau 4.7. Probabilités de bonne détection et fausse alarme des signaux de la base REELSIM avant et après application de l'algorithme du choix de la meilleure base de détection.

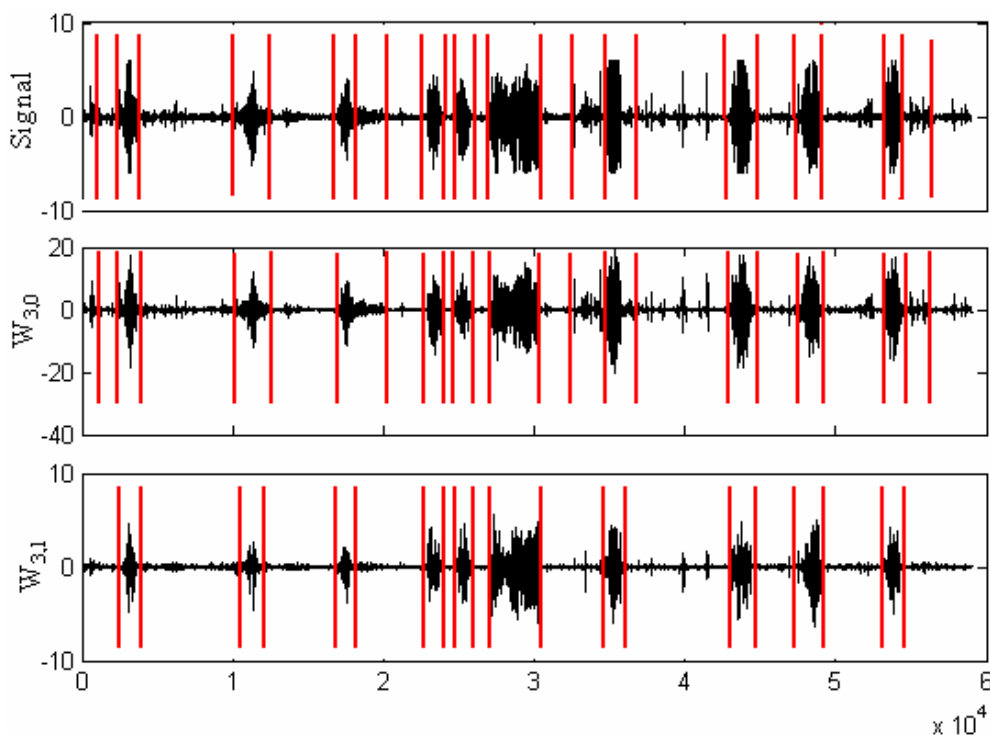


Figure 4.16. Exemple de signal EMG utérin. L'algorithme de détection est appliqué sur les paquets 7 et 8. Les instants de détection du signal sont obtenus après correction et fusion.

4.6 Performance liée à l'utilisation directe des coefficients d'ondelettes

Pour montrer la performance de l'emploi des coefficients des paquets d'ondelettes au lieu des signaux reconstruits dans le cas de la détermination des paquets les plus pertinents pour la détection, nous avons calculé la statistique de K-S résultant de l'estimation de la distance de K-L (équation 3.13) pour chaque paquet en utilisant les signaux reconstruits. Nous avons utilisé le même ensemble d'apprentissage G1SIM1 de la base SIM1 pour lequel les moyennes des statistiques de K-S avaient été présentées par ordre décroissant sur la figure 3.10 (cas des coefficients des paquets d'ondelettes). La figure 4.17 montre que l'ordre des paquets est conservé dans le cas de l'utilisation des signaux reconstruits. Les mêmes paquets représentant des ruptures ont été obtenus dans le cas de l'utilisation des coefficients des paquets d'ondelettes et dans le cas des signaux reconstruits. Par contre le seuil à utiliser pour la sélection des paquets n'est pas aussi clair que dans le cas d'utilisation directe des coefficients (figure 3.10) ce qui conforte encore notre choix de travail direct sur les coefficients. Pour tester la réduction de temps de calcul obtenu en utilisant les coefficients d'ondelettes, nous avons estimé d'une part le temps de calcul de l'algorithme de détection DCS appliqué directement sur les coefficients d'ondelettes, d'autre part le temps de calcul de DCS appliqué sur les signaux reconstruits.

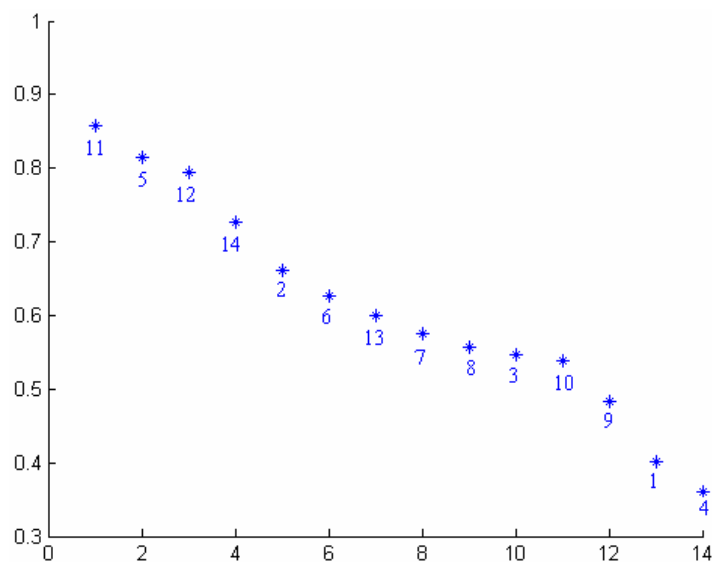


Figure 4.17. Moyennes des valeurs de la statistique D_{\max} de K-S obtenues pour les 14 paquets à partir de 1000 signaux simulés de la base SIM1 dans le cas des signaux reconstruits. Axe X : unité arbitraire. Axe Y: Valeurs de D_{\max} .

La figure 4.18 confirme bien que l'emploi direct des coefficients d'ondelettes est plus rapide que l'emploi des signaux reconstruits [Che05a], [Che05c]. Plus la longueur du signal augmente, plus l'emploi direct des coefficients sera efficace du point de vue du temps de calcul.

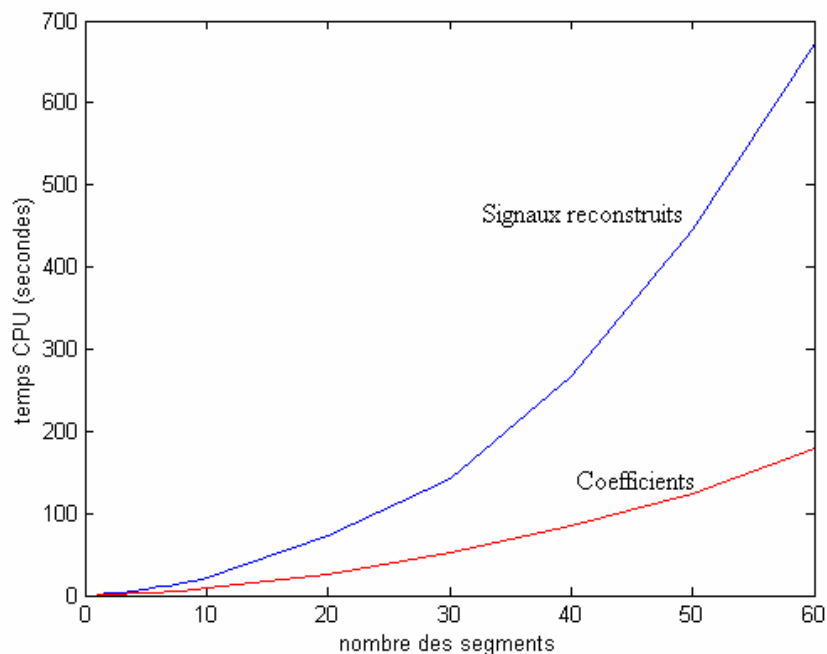


Figure 4.18. Comparaison entre les temps de calcul de l'algorithme de détection DCS dans les deux cas de l'emploi des coefficients d'ondelettes et de leurs signaux reconstruits.

4.7 Détection en temps réel

L'un des objectifs de ces travaux de recherche est la détection en temps réel, c'est-à-dire le fait de pouvoir effectuer la détection en même temps que l'acquisition des signaux, en ne tolérant qu'un retard sans cumul indispensable à l'estimation et celui imposé par la détection. Dans le cadre de la décomposition en ondelettes ou paquets d'ondelettes, à chaque acquisition d'un point du signal, il est tout à fait possible d'effectuer le calcul de nouveaux coefficients sur les différentes échelles choisies ou sur les différents paquets choisis, puis d'utiliser ces nouveaux coefficients en ligne pour les algorithmes de détection.

Les étapes de détection en temps réel se résument comme suit :

- Extension des coefficients de chaque paquet pour assurer la convolution avec le filtre en cours (dépendant de l'ondelette utilisée), jusqu'à obtention du nombre de points suffisant pour la convolution. Ceci se fait pendant la phase d'initialisation au début de l'acquisition des échantillons.

- La détection s'effectue sur les paquets déjà choisis dans une phase d'apprentissage par application de l'algorithme du choix de la meilleure base. De ce fait, le calcul des coefficients est effectué pour l'ensemble des paquets amenant aux paquets choisis.
- L'algorithme de détection retenu (DCS) nécessite l'existence de N points avant et après l'instant en cours. Pour une détection en temps réel, la procédure de détection sur un paquet donné commence quand $2.N$ coefficients ont déjà été calculés, et la détection ne se poursuit qu'après le calcul du nouveau coefficient, pour avoir toujours N points avant et après l'instant en cours.
- Chaque instant de détection est corrigé.
- L'algorithme de fusion est appliqué sur les intervalles de détection superposés.

Le nombre de points initiaux nécessaires à cette modélisation est cependant supposé déjà disponible. Pour la décomposition, ce nombre est égal à la longueur du filtre passe bas pour le calcul des coefficients d'approximation, et égal à la longueur de filtre passe haut pour le calcul des coefficients de détail dans le cas de DWT. Il est, enfin, égal à la longueur du filtre utilisé pour le calcul des coefficients de chaque paquet dans le cas de WPT.

Pour la détection en temps réel, supposons qu'on désire détecter les événements sur un paquet $W_{j,n}$. Alors il est nécessaire de calculer tous les coefficients des paquets amenant au paquet $W_{j,n}$ en commençant par le signal original. Par exemple pour appliquer l'algorithme de détection sur le paquet 7 ($W_{3,0}$) alors il est nécessaire de calculer les coefficients des paquets 1 et 3 ($W_{1,0}$ et $W_{2,0}$).

Chaque coefficient est le résultat d'une convolution avec le niveau précédent. Le calcul d'un coefficient d'un paquet d'ondelettes de niveau j et d'index $2n$ ou $2n + 1$ nécessite l'emploi de L coefficients du paquet précédent de niveau $j - 1$ et d'index n (L étant le nombre de points de la réponse impulsionnelle des filtres correspondant à l'ondelette utilisée).

La décimation consiste à éliminer un coefficient sur deux. Le calcul des coefficients sur chaque paquet ne se fait alors que si l'indice du coefficient du paquet à l'échelle précédente est pair.

Comme le calcul des coefficients des paquets d'ondelettes est récursif, ceci aide à calculer le coefficient du paquet $W_{j,n}$ dès que les paquets auxquels il est lié aux échelles précédentes ont le nombre suffisant de coefficients pour permettre le calcul du nouveau coefficient au niveau du paquet $W_{j,n}$. Rappelons les équations du calcul des coefficients des paquets d'ondelettes :

$$C_{j,2n}(k) = \sum_l \hat{h}(2k-l)C_{j-1,n}$$

$$C_{j,2n+1}(k) = \sum_l \hat{g}(2k-l)C_{j-1,n}$$

Ainsi le paquet $C_{j,2n}$ (respectivement $C_{j,2n+1}$) est obtenu par filtrage passe-bas (resp. passe haut) de $C_{j-1,n}$ par le filtre g (resp. h).

La décomposition discrète en ondelettes étant un cas particulier de la décomposition en paquets d'ondelettes, son approche temps réel est la même que pour la WPT.

En ce qui concerne la détection, si l'algorithme DCS est appliqué sur le paquet $W_{j,n}$, il est nécessaire de connaître N points avant et N points après l'instant courant k . En traitement "temps" réel, un retard à la détection est nécessairement induit par le calcul de ces N points après l'instant courant nécessaire à l'estimation des paramètres des distributions "avant" et "après", et au calcul des fonctions (4.25 et 4.26). A chaque fois qu'un instant de changement est détecté, il est corrigé selon l'équation (4.9). Si la détection se fait sur plusieurs paquets, et que des instants de changement sont détectés sur des intervalles superposés, l'algorithme de fusion (paragraphe 4.4.3) s'applique à ces intervalles de détection superposés pour obtenir un instant de changement appliqué au signal original pour la séquence déjà acquise, sans attendre l'acquisition complète du signal.

Pour illustrer la performance de la décomposition en paquets d'ondelettes et la détection en temps réel, sans attendre l'acquisition complète du signal en cours, les enregistrements réels de longue durée des signaux femmes de la base REEL2 ont été utilisés. La détection a été faite en introduisant point par point les échantillons du signal femme. La performance a été évaluée en calculant la moyenne et l'écart type des différences entre les instants de changements estimés par notre algorithme et par ceux indiqués par l'expert. Nous avons également relevé le taux de non détection qui sont les instants de changement non détecté par notre algorithme alors qu'ils avaient été relevés par l'expert. Le tableau 4.8 indique les moyennes et les écarts types des différences relevées entre algorithme de détection et expert. Les résultats sont donnés en nombre de points et en secondes. Le retard est en moyenne de 1.07 seconde pour un écart type de 4.76. L'ondelette Symlet 5 a été utilisée (longueur du

	Moyenne	Ecart type
Nombre de points	34.17	152.19
secondes	1.07	4.76

Tableau 4.8. Moyennes et écart types de la différence entre les instants de changement estimés et ceux précisés par l'expert.

filtre 10) car comme nous l'avons vu c'est l'ondelette qui a donné les meilleurs résultats au niveau de la correction des instants de changement. Malgré cette propriété importante, d'autres ondelettes auraient pu être utilisées sur la base de leurs performances pour la détection. Cependant, compte tenu des résultats très encourageants obtenus avec l'ondelette Symlet 5, nous n'avons pas jugé utile d'aller plus avant dans le choix éventuel d'une autre ondelette.

D'autre part, il est difficile de faire la part des choses entre les performances liées à l'ondelette choisie et celles induites par l'algorithme de détection lui-même.

L'algorithme de somme cumulée dynamique (DCS) développé par Khalil [Kha99] a été utilisé dans notre cas au lieu de celui proposé ultérieurement par El Falou [Fal02], pour des raisons de simplicité de mise en œuvre en temps réel. Cependant, soulignons que le fait de choisir d'autres algorithmes séquentiels est complètement indépendant de la méthodologie adoptée pour le choix de la meilleure base et ne remet nullement en cause les performances associées.

Les méthodes de correction et de fusion sont appliquées en ligne chaque fois qu'un instant de détection est obtenu. Le fait de fusionner les intervalles superposés et le choix du calcul du barycentre peuvent également dans une certaine mesure modifier la position exacte de l'instant de changement par rapport au signal original.

N'oublions cependant pas que l'avis de l'expert peut lui-même comporter des erreurs. L'absence d'un nombre suffisant d'experts ne nous a pas permis de faire des études de concordance, mais la présentation du même enregistrement à plusieurs reprises au même expert a révélé un certain nombre de discordances.

Le pourcentage de non détection est quant à lui de 9.519 %. Ce résultat est obtenu en comptant le nombre total de changements identifiés par l'expert et les instants non détectés par notre algorithme.

Le pourcentage du nombre de fausses alarmes est quant à lui de 10%. Ce résultat est obtenu en comptant les instants détectés par notre algorithme et non indiqués par l'expert.

Parfois seuls les débuts d'événements très courts comme les mouvements actifs fœtaux sont détectés alors que leurs fins ne sont pas détectés. De plus le choix du seuil est un élément critique de la méthode DCS, alors que dans sa version modifiée (MDCS), le système de double seuil et d'autres apprentissages de ces seuils contribuent à rendre la dépendance au seuil beaucoup moins critique.

4.8 Conclusion

Partant du fait que toute l'information est contenue dans les coefficients des paquets d'ondelettes, le procédé de détection a été appliqué aux coefficients plutôt qu'aux signaux reconstruits.

L'inconvénient principal de l'utilisation directe des coefficients est le décalage entre les vrais instants de changement et ceux détectés sur les échelles ou paquets d'ondelettes. Ceci est dû à la décimation et à l'effet de convolution de l'ondelette utilisée.

Dans ce chapitre, nous n'avons pas développé de façon extensive la description des différents algorithmes de détection, partant du principe que la comparaison des performances effectuée ici était valide avant et après la réduction du nombre de paquets d'ondelettes quelque soit l'algorithme utilisé.

Nous avons en fait présenté l'algorithme de somme cumulée dynamique DCS parce qu'il a été développé auparavant pour la segmentation de l'EMG utérin, qu'il avait alors montré son efficacité [Kha99b], et qu'il était de mise en œuvre plus simple que sa version modifiée (MDCS).

Dans le but d'appliquer l'algorithme de détection sur les coefficients de détail ou sur les coefficients des meilleurs paquets choisis par l'algorithme de sélection de la meilleure base (et non sur les signaux reconstruits), un algorithme de correction des instants de changements a été mis en œuvre. Les instants de changements sont recalés de façon à ce qu'ils correspondent à l'échelle temporelle du signal original. La performance de la correction proposée dépend de l'ondelette utilisée pour la transformation et de l'influence des coefficients à l'échelle précédente. Les résultats sont très satisfaisants pour ce qui concerne la moyenne du temps corrigé (entre 4 et 6 points) mais ils varient considérablement en ce qui concerne l'écart type selon l'ondelette choisie. L'approche barycentrique, bien que la position du barycentre ne soit pas en général une valeur entière, nous a semblé une bonne approche pour tenir compte de la spécificité de chaque ondelette. D'autres approches, comme par exemple la valeur maximale

de la réponse impulsionnelle du filtre, pourraient être également envisagées. Cependant les résultats produits par l'approche barycentrique étaient tout à fait acceptables, en particulier pour l'ondelette Symlet 5.

D'autre part, ce chapitre a montré la nécessité de fusion de différents instants de détection identifiés sur les niveaux de détails ou sur les paquets sélectionnés pour obtenir un seul instant de changement pour une rupture donnée. Un algorithme de fusion a été développé pour fournir, pour un ensemble d'intervalles de détection superposés, un seul instant de changement rapporté au signal original. Ces procédures ont été appliquées sur des signaux simulés et sur l'EMG utérin.

La détection en temps réel a été effectuée sur les enregistrements réels de longue durée disponibles. Le choix de l'ondelette, l'avis de l'expert et le seuil de l'algorithme de détection ont leur importance dans la qualité des résultats. Pour les paramètres que nous avons choisis, de bons résultats ont été obtenus.

Après obtention des différents segments contenus dans l'EMG utérin, le but est de faire correspondre ces segments à des phénomènes physiologiques connus *a priori* comme les contractions, les mouvements actifs foetaux, les ondes d'Alvarez, etc. Cette étape finale est développée dans le chapitre suivant, qui propose une procédure du choix des paquets les plus discriminants pour les enregistrements d'EMG utérin dans le cas de décomposition en paquets d'ondelettes. La mise en oeuvre d'une méthode de classification des événements en lien avec les phénomènes physiologiques qu'ils représentent, sera également présentée.

Chapitre 5

Classification et identification des événements de l'EMG utérin

5.1 Introduction

Les événements segmentés de l'EMG utérin ont été obtenus après l'application d'un algorithme de détection sur les meilleurs paquets sélectionnés pour cet objectif de détection. Rappelons que les 4 types d'événements sont : les contractions, les ondes Alvarez, les mouvements actifs fœtaux, et les ondes LDBF (Longue Durée Basse Fréquence). Ils forment 4 classes. Le bruit réel constitué des segments inter - événements forme la cinquième classe.

Chaque événement est caractérisé principalement par son contenu fréquentiel et sa longueur (voir chapitre 1). L'approche proposée pour la réalisation d'un classifieur s'adresse de nouveau à la décomposition en paquets d'ondelettes. C'est une approche sans modèle utilisant la connaissance d'un ensemble d'apprentissage constitué par expertise. Le principe mis en œuvre consiste à décomposer des événements de référence sur un arbre complet de paquets d'ondelettes puis à choisir parmi tous les paquets ceux qui permettent de discriminer le mieux entre les différents événements. Tout nouvel événement, obtenu par la procédure de segmentation décrite dans les chapitres précédents, sera alors décomposé uniquement sur les paquets sélectionnés. Ce seront en particulier les propriétés des coefficients de ces paquets qui serviront de base au classement de l'événement analysé.

Une question très importante se pose en effet en classification, et, plus généralement, dans tout problème de modélisation statistique : celle du choix des caractéristiques discriminantes (dont la connaissance est susceptible de contribuer utilement à la solution du problème posé) parmi un ensemble de caractéristiques possibles.

Nous avons déjà souligné que la transformée en paquets d'ondelettes était utilisée dans de nombreux domaines comme le débruitage, la compression [Wic91], [Wal97] et la reconnaissance de formes [Coi92], [Pit99], [Sai95], [Sai02], [Sai97]. On trouve également des méthodes de classification basées sur la transformée en paquets d'ondelettes qui utilisent des caractéristiques telles que l'énergie [Lea95] ou la distance [She97].

Dans [Tur03] un système intelligent est présenté pour l'interprétation du signal Doppler pour les maladies cardiaques, méthode fondée sur la reconnaissance de modèle. Ce travail combine l'extraction des caractéristiques en utilisant la décomposition en paquets d'ondelettes et la classification en utilisant un réseau de neurones de type perceptron multicouches. Kalayci [Kal95] a choisi certains coefficients de chaque niveau de décomposition pour qu'ils servent d'entrées à un réseau de neurones pour la détection des transitoires dans les électroencéphalogrammes. [Yan05] a proposé une approche pour le diagnostic en santé, basé sur la décomposition en paquets d'ondelettes. Les caractéristiques extraites des paquets d'ondelettes ont été choisies dans ce cas selon le critère de discrimination de Fisher. Elles ont été utilisées comme entrées d'un classifieur à base de réseau de neurones pour décider de l'état du système.

[Akh05] a utilisé la transformée en ondelettes pour extraire les caractéristiques essentielles du ballistocardiogramme (BCG) et les réseaux de neurones artificiels pour classifier les différents sujets traités. [Ube04] a proposé des méthodes d'extraction des caractéristiques utilisées pour le diagnostic automatique des maladies artérielles. Quelques conclusions ont été tirées montrant l'efficacité de la décomposition en ondelettes pour l'extraction des caractéristiques utilisées ensuite pour le diagnostic de maladies artérielles carotides ophtalmiques et internes. Dans [Sun05], une conjonction entre la décomposition discrète en ondelettes et les réseaux de neurones probabilistes (PNN Probabilistic Neural Network), a été proposée pour classifier les tissus à partir de données génétiques. Dans [Xu02] des caractéristiques statistiques ont été extraites des coefficients d'ondelettes et un algorithme génétique a été utilisé pour sélectionner les caractéristiques optimales. Les coefficients d'ondelettes obtenus à partir de signaux ultrasonores ont été considérés comme caractéristiques d'entrée de deux classifieurs : réseaux de neurones artificiels et Machines à Vecteurs Supports (SVM) [Lee05]. Des caractéristiques d'images ont été extraites après

décomposition discrète en ondelettes et utilisées comme des entrées d'un classifieur SVM pour la classification de la texture [Raj04]. Une fonction fondée sur le calcul de la distance de Kullback Leibler a été utilisée pour choisir les paquets les plus discriminants [Raj02]. Dans [She01] et [Yon05], les signaux EEG ont été classifiés en utilisant d'abord la décomposition en paquets d'ondelettes, puis un classifieur SVM.

Dans le présent travail, la même approche générale a été mise en œuvre :

- sélection des meilleurs paquets d'une décomposition en paquets d'ondelettes, au sens de la discrimination.
- extraction des caractéristiques des événements à partir des paquets sélectionnés.
- Classement des nouveaux événements par une méthode de classification (les performances de quelques méthodes seront présentées dans la suite du chapitre).

Les caractéristiques discriminantes seront calculées ici directement à partir des coefficients des paquets d'ondelettes (Energie). Une caractéristique supplémentaire, la durée des événements, sera prise en considération également. Des études antérieures ont montré en effet son importance en terme de discrimination [Kha99]. La sélection des caractéristiques les plus discriminantes se fera en utilisant un critère de discrimination basé sur le calcul du rapport entre variance intra classes et variance totale. Les caractéristiques choisies seront introduites dans plusieurs classifieurs tels que la méthode du K plus proche voisin [Lea92], la distance de Mahalanobis [McI92], les réseaux de neurones [Kal95], et les Machines à Vecteurs Supports [Yon05], etc.

5.2 Reconnaissance des Formes et paquets d'ondelettes

La Reconnaissance des Formes (RdF) constitue l'apprentissage ou la découverte de structures appelées classes dans un ensemble de données généralement contaminées par du bruit.

Suivant le domaine d'application visé, l'information traitée est de nature qualitative, quantitative ou les deux à la fois. Dans le cas des événements EMG, l'intérêt de la reconnaissance des formes revient à savoir à quelle classe physiologique appartient un segment donné.

L'objectif de la reconnaissance des formes est de classifier des événements en classes à partir d'observations effectuées sur ceux-ci. Dans le cas d'utilisation de la décomposition en paquets d'ondelettes, la méthodologie peut se décliner en 4 étapes:

1. Acquisition des signaux.
2. Segmentation (incluant déjà dans notre cas une décomposition en paquets d'ondelettes).
3. Décomposition en paquets d'ondelettes pour la classification.
4. 4.1 Sélection de la meilleure base pour la discrimination (phase d'apprentissage).
4.2 Classification (phase de routine).

Dans la phase d'apprentissage, nous adopterons une démarche supervisée, constituée d'événements étiquetés. Comme pour le problème précédent de détection, il s'agira alors de rechercher, dans une décomposition en paquets d'ondelettes, les paquets les plus pertinents pour identifier correctement ces événements.

Pour pouvoir aborder le problème de la sélection des meilleurs paquets, il est nécessaire de définir :

- les paramètres représentatifs des paquets.
- le critère de base à la sélection des paquets les plus discriminants.

Les paquets ont été évoqués au paragraphe précédent. Compte tenu de l'effet filtré de la WPT, et du fait que le contenu fréquentiel est un élément majeur de discrimination, un seul paramètre, l'énergie représentée par les coefficients d'un paquet, sera extrait de la WPT. Un paramètre additionnel, la durée de l'événement, sera introduit plus tard dans le processus final de classification.

Le choix du critère à utiliser pour sélectionner la meilleure base en vue d'une classification est par contre discuté au paragraphe suivant.

5.3 Sélection de la meilleure base pour la discrimination entre classes

5.3.1 Critère de discrimination entre classes pour la sélection de la meilleure base

Le critère de classification permettant de sélectionner les meilleurs paquets dans une optique de classification doit naturellement être différent de celui utilisé en détection ou en compression. Plusieurs algorithmes du choix de la meilleure base ont été mis en oeuvre pour la classification. Saito [Sai94] a proposé un algorithme spécifique pour ce choix de la

meilleure base en classification (Local Discriminant Best Bases (LDB)). Dans [Sai97], Saito a introduit l'information lithologique des ondes entières pour un problème de classification, i. e. classifier les types des roches en utilisant la forme des ondes. Il a appliqué la méthode du choix de la meilleure base discriminante (LDB) pour différencier entre les événements.

Dans [Sai94] Saito a développé tout d'abord l'algorithme du choix de la meilleure base pour la classification en étendant celui de Coifman et Wicherhauser utilisé pour la compression. Il a utilisé l'entropie relative (aussi connue comme cross entropy, distance de Kullback-Leibler, ou I-divergence), déjà utilisée par Coifman [Coi92], comme une fonction de coût pour mesurer la discrimination entre les différentes caractéristiques. Par exemple, l'entropie relative définie comme l'information de discrimination $D(\mathbf{x}, \mathbf{y})$ entre deux séquences normalisées $\{x_i\}$ et $\{y_i\}$ est donnée par :

$$I(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i \log \frac{x_i}{y_i} \quad (5.1)$$

Pour compenser le problème de dissymétrie, il définit :

$$D(\mathbf{x}, \mathbf{y}) = I(\mathbf{x}, \mathbf{y}) + I(\mathbf{y}, \mathbf{x}) \quad (5.2)$$

Dans notre cas, les séquences $\{x_i\}$ et $\{y_i\}$ représentent les séquences de coefficients obtenus sur chacun des paquets d'une décomposition en paquets d'ondelettes.

Pour mesurer l'information de discrimination entre plusieurs classes, l'équation 5.2 devient :

$$D\left(\left\{x^{(c)}\right\}_{c=1}^M\right) = \sum_{i=1}^{M-1} \sum_{j=i+1}^M D(x^{(i)}, x^{(j)}) \quad (5.3)$$

où M est le nombre total des classes. Il est clair que l'équation (5.3) est additive.

Plusieurs faiblesses intrinsèques existent dans la formule (5.3), parmi lesquelles le fait qu'un grand nombre des coefficients des paquets d'ondelettes d'une séquence $\{x_i\}$ sont petits ou nuls, ce qui mène, par l'équation (5.1) à des valeurs très grandes de l'équation (5.3).

Par conséquent, l'entropie relative ne peut pas être utilisée d'une façon efficace comme critère pour le problème de sélection de la meilleure base dans notre cas.

Le critère à retenir pour mesurer l'efficacité d'une décomposition spécifique en paquets d'ondelettes doit néanmoins être une mesure de la capacité des paquets retenus à séparer au mieux les classes. Un des critères est alors de minimiser la variance intra classe et de

maximiser la variance inter-classe. [She97] a utilisé un critère de séparabilité qui minimise la distance euclidienne intra classe et qui maximise la distance euclidienne entre les classes. [Fuk90] a utilisé ce critère pour sélectionner les coefficients des paquets d'ondelettes les plus discriminants. C'est donc le rapport variance intra classe sur variance totale que nous adopterons dans la suite.

Calculer des variances (intra classe, inter classe ou totale) nécessite de se doter de paramètres extraits des séquences de coefficients des paquets, qui soient les plus pertinents en matière de discrimination.

Pour ce qui concerne les événements de l'EMG, comme nous l'avons indiqué dans les chapitres précédents, ils se caractérisent essentiellement par leur contenu fréquentiel. Le calcul de la variance (de l'énergie) des coefficients au niveau de chaque paquet nous a donc paru une approche raisonnable pour caractériser les événements.

La démarche d'apprentissage revient donc à décomposer un ensemble d'apprentissage étiqueté sur un arbre de décomposition complet, puis de calculer, pour chaque paquet l'énergie moyenne de ses coefficients, et enfin de déterminer, toujours pour chaque paquet, le rapport entre variance intra classe et variance totale.

Pour un niveau de décomposition J, le nombre N des paquets à considérer est :

$$N = 2^{J+1} - 1 \quad (5.4)$$

Soit M le nombre total de classes. La $i^{\text{ème}}$ classe est composée de m_i éléments. Le centre de gravité \hat{g}_i de la $i^{\text{ème}}$ classe est défini par :

$$\hat{g}_i^n = \frac{1}{m_i} \sum_{k=1}^{n_i} x_{ik}^n \quad (5.5)$$

où x_{ik}^n est le $k^{\text{ième}}$ élément de la classe i ($i = 1..5$), et du paquet numéro n .

Soit \hat{g}^n , le centre de gravité global du paquet n :

$$\hat{g}^n = \frac{1}{m} \sum_{k=1}^M m_k \hat{g}_k^n \quad (5.6)$$

où m est le nombre total des échantillons.

La variance intraclasse $\hat{\Sigma}_w^n$ d'un paquet n est définie par :

$$\hat{\Sigma}_w^n = \frac{1}{m} \sum_{i=1}^M \sum_{k=1}^{m_i} \left(x_{ik}^n - \hat{g}_i^n \right)^2 \quad (5.7)$$

La variance inter - classes au niveau d'un paquet n $\hat{\Sigma}_B^n$ s'écrit :

$$\hat{\Sigma}_B^n = \frac{1}{m} \sum_{i=1}^M m_i \left(\hat{g}_i^n - \hat{g}^n \right)^2 \quad (5.8)$$

Il est facile de démontrer que la variance totale $\hat{\Sigma}^n$ est égale à la somme des variances interclasse et intra classe [Dub90]:

$$\hat{\Sigma}^n = \hat{\Sigma}_w^n + \hat{\Sigma}_B^n \quad (5.9)$$

Pour chaque paquet n , définissons le critère de discrimination comme suit :

$$R^n = \frac{\hat{\Sigma}_w^n}{\hat{\Sigma}^n} \quad (5.10)$$

Le rapport R^n (équation 5.10) calcule le rapport entre la variance intra classe et la variance totale. Si les distances entre les classes sont importantes et chaque classe fortement concentrée, alors le rapport R^n du paquet n est petit. Dans ce cas, les classes sont bien séparées. Ceci indique que le paquet n est parmi les paquets ayant une bonne qualité discriminante.

Il reste à vérifier que le rapport R^n est un bon critère de discrimination entre les classes d'événements dans le cas de l'EMG utérin lorsque ceux-ci sont caractérisés par la variance des coefficients des paquets d'ondelettes.

5.3.2 Sélection de la meilleure base pour la classification

L'objectif est donc ici de ne retenir que les paquets qui possèdent un caractère discriminant. La caractérisation d'un paquet n se fait de la façon suivante : si la valeur du rapport R^n est au dessous d'un seuil donné, la valeur "1" est associée à ce paquet. Dans le cas contraire, la valeur "0" y est associée.

Comme l'arbre de décomposition en paquets d'ondelettes est formé de paquets redondants, l'étape suivante consiste, en général, à réduire le nombre des paquets sélectionnés à ceux qui portent des informations complémentaires.

La première phase de la réduction de redondance (sélection de la meilleure base) suit en tous points l'algorithme défini dans le cadre de la sélection de la meilleure base pour la détection.

Cependant, ici, nous pouvons compléter ce processus de sélection de la meilleure base en utilisant de la connaissance *a priori* sur les bandes de fréquence des divers événements contenus dans l'EMG utérin.

L'idée est alors de balayer les paquets éliminés dans la phase précédente de réduction de redondance, et de réintégrer ceux qui correspondent à des bandes de fréquence très corrélées à celles des événements. Cette démarche heuristique complémentaire ne remet en aucun cas en cause le processus de sélection des paquets. Elle tend simplement à limiter la suppression des paquets redondants. Nous verrons en effet dans la suite du chapitre que, contrairement à la sélection de la meilleure base dans le cas de la détection, le même processus de réduction de redondance appliqué à la classification, peut dans certains cas mener à des résultats dégradés.

La figure 5.1 présente un exemple de sélection de la meilleure base dans le cas de classification.

Dans le processus de sélection des nœuds discriminants, le fait de conserver à chaque niveau le père ou ses fils n'a pas été garanti. La sélection finale ne constitue pas généralement une base complète. Cette restriction ne présente cependant pas de difficultés puisque l'objectif n'est pas de reconstruire un signal mais uniquement de fournir des informations pertinentes à une classification.

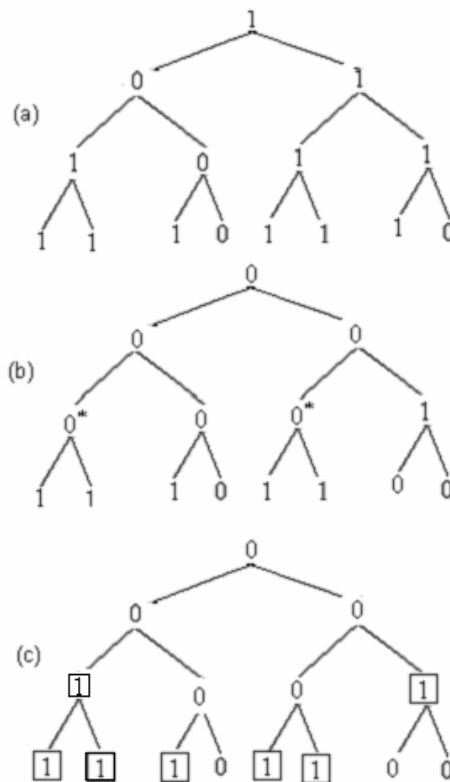


Figure 5.1. Etapes de construction de la meilleure base pour la classification : a- Mise à "1" des paquets ayant les plus petites valeurs de R^n et à "0" les autres. b- Modification des valeurs des nœuds (réduction de redondance). c- Requalification des paquets ayant un contenu fréquentiel très proche de celui de certains types d'événements. La meilleure base est constituée des nœuds à 1.

5.4 Méthodes de classification

A partir d'un ensemble d'individus déjà classés par l'expert, le système peut apprendre à classer comme l'expert. Après apprentissage, le système est capable de classer de nouveaux individus. Il existe de nombreuses techniques de classification qui peuvent être employées pour développer une règle de la décision. Plusieurs de ces techniques sont décrites très succinctement dans les paragraphes suivants, et leurs performances sont ensuite comparées dans le cadre de l'application à l'EMG utérin.

5.4.1 Règle du plus proche voisin

La méthode du plus proche voisin se caractérise par sa capacité à traiter localement les informations. Elle consiste à examiner l'élément, dont la classe est connue, le plus proche de l'élément dont on veut déterminer la classe.

Soit $X^n = \{\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_n\}$ un ensemble d'apprentissage formé de n vecteurs indépendants, où chaque événement est représenté par un vecteur paramétrique, \mathbf{x}_j . La classe de chaque élément de l'ensemble d'apprentissage est connue, elle sera désignée pour l'élément \mathbf{x}_j par $w(\mathbf{x}_j)$. L'événement à classer, représenté par un vecteur paramétrique \mathbf{x} , est affecté à la classe correspondant au voisin le plus proche parmi l'ensemble d'apprentissage. La règle dite de 1PPV est simplement la suivante :

$$\hat{w}(\mathbf{x}) = w(\mathbf{x}_{PPV}) \quad \text{si} \quad d(\mathbf{x}, \mathbf{x}_{PPV}) = \underset{j=1..n}{\text{Min}} \quad d(\mathbf{x}, \mathbf{x}_j) \quad (5.11)$$

où \mathbf{x}_{PPV} est l'échantillon le plus proche de \mathbf{x} , et $\hat{w}(\mathbf{x})$ est la classe d'affectation estimée de \mathbf{x} .

La règle du 1PPV classe \mathbf{x} selon sa classe d'affectation estimée $\hat{w}(\mathbf{x})$. La notion du "plus proche", en termes des distances les plus petites, laisse un certain choix pour la détermination de la distance. Dans ce travail, la distance euclidienne sera utilisée.

Dans le cas des K plus proches voisins, pour chaque vecteur \mathbf{x} à classer, on recherche les K plus proches voisins dans l'échantillon d'apprentissage, et \mathbf{x} est affecté au groupe majoritaire [Gov03].

5.4.2 Distance de Mahalanobis et classification

La distance de Mahalanobis est l'une des mesures capables de discriminer ou de séparer entre les classes [McI92]. C'est une méthode globale linéaire pondérée par le volume de la classe en prenant en compte l'ensemble des échantillons d'une classe. Soit X_i^n l'ensemble des échantillons de l'ensemble d'apprentissage X^n appartenant à la classe i .

Soient $\mu = E(X_i^n)$ et $\Sigma(X_i^n)$ l'espérance et la matrice de variance-covariance de l'ensemble des échantillons de cette classe.

La distance de Mahalanobis D_i utilisée dans notre cas est la distance entre l'ensemble d'apprentissage pour la classe i et un élément \mathbf{x} dont on désire déterminer la classe.

$$D_i^2 = (\mathbf{x} - \mu_i)' \Sigma_i^{-1} (\mathbf{x} - \mu_i) \quad (5.12)$$

D_i^2 est appelée distance de Mahalanobis de \mathbf{x} à μ_i .

La classe d'affectation de \mathbf{x} sera celle qui produira la distance minimale :

$$\hat{w}(\mathbf{x}) = \text{card}(\min_i D_i) \quad (5.13)$$

5.4.3 Réseaux de Neurones : Perceptron multicouches

Les réseaux de neurones possèdent la propriété d'approcher d'une façon parcimonieuse n'importe quelle fonction bornée (par exemple une probabilité). Les poids des connexions dans les réseaux de neurones sont généralement estimés par minimisation d'un critère de moindres carrés, qui permet d'estimer les probabilités *a posteriori*.

Les réseaux Perceptron Multicouche (Multi Layer Perceptron : MLP) ont été utilisés pour résoudre des problèmes difficiles par apprentissage supervisé à partir d'un algorithme de "retro-propagation de l'erreur" qui se base sur la règle d'apprentissage par correction d'erreur. Le processus de rétropropagation d'erreur consiste à effectuer deux passages à travers les différentes couches du réseau: un passage direct et un passage rétrograde.

Pendant le passage direct, les poids synaptiques du réseau sont fixes, et en rétrograde les poids synaptiques sont ajustés à partir d'une règle de correction de l'erreur. Ces ajustements continuent jusqu'à ce que la réponse du réseau s'approche de la réponse désirée au sens de l'ensemble d'apprentissage ou de la fonction.

1. Architecture du réseau

Considérons un Perceptron Multicouches formé d'une couche de L entrées, d'une couche de S neurones cachés et d'une couche de M neurones de sortie. Le réseau est entièrement connecté, c.a.d. chaque neurone d'une couche est relié à l'ensemble des neurones de la couche suivante.

Entrées : $x_n : n = 0, 1, \dots, L ; x_0 = 1$.

Sorties : $y_k : k = 1, \dots, M$.

Sorties internes : $z_j : j = 1, \dots, S ; z_0 = 1$.

Equations de fonctionnement :

Connexions entrées- neurones cachés :

$$u_j = \sum_{n=0}^L w_{jn} x_n ; j = 1, \dots, S \quad (5.14)$$

Fonctions d'activation :

$$z_j = \Phi(u_j) \quad (5.15)$$

Connexions neurones cachés-sorties:

$$v_k = \sum_{j=0}^S w_{kj} z_j ; k = 1, \dots, M \quad (5.16)$$

Fonction d'activation :

$$y_k = \Phi(v_k) \quad (5.17)$$

2. Stratégie d'apprentissage

L'erreur commise par le réseau, définie comme la différence entre la sortie attendue et la sortie du réseau, va permettre d'adapter les poids de la structure afin de minimiser cette erreur. Plusieurs algorithmes existent afin de réaliser cet apprentissage. Nous pouvons citer par exemple la rétropropagation du gradient [Ros62].

L'erreur au carré entre les sorties désirées et les sorties calculées par le réseau est :

$$E = \frac{1}{2} \sum_{k=1}^M (d_k - y_k)^2 \quad (5.18)$$

5.4.4 Machine à Vecteurs Supports

Les "Support Vector Machines" (SVM) sont de nouvelles techniques discriminantes dans la théorie de l'apprentissage statistique. Elles ont été proposées en 1995 par V.Vapnik [Vap95]. Elles permettent d'aborder des problèmes divers comme la régression, la classification, la fusion etc. La méthode SVM est basée sur la construction de l'hyperplan séparateur "optimal". Cet hyperplan optimal est construit de manière à maximiser la distance minimale de l'hyperplan aux exemples d'apprentissage.

Deux cas se présentent pour la construction de l'hyperplan optimal séparant des données appartenant à deux classes différentes : le cas des données linéairement séparables et le cas des données non-linéairement séparables [Kha02].

Soit l'ensemble D tel que :

$$D = \left\{ (x_i, y_i) \in \mathfrak{R}^n \times \{-1, 1\} \text{ pour } i = 1, \dots, m \right\} \quad (5.19)$$

a. Cas des données linéairement séparables

Dans ce paragraphe, nous présentons la méthode générale de construction de l'Hyperplan Optimal (HO) qui sépare des données appartenant à deux classes différentes linéairement séparables.

Soit $H : (w.x) + b$ l'hyperplan qui satisfait les conditions suivantes :

$$\begin{cases} w.x_i + b \geq 1 & \text{si } y_i = 1 \\ w.x_i + b \leq -1 & \text{si } y_i = -1 \end{cases} \quad (5.20)$$

Ce qui est équivalent à :

$$y_i(w.x_i + b) \geq 1 \text{ pour } i = 1, \dots, m \quad (5.21)$$

Un hyperplan optimal (HO) est un hyperplan qui maximise la marge O qui représente la plus petite distance entre les différentes données des deux classes et l'hyperplan. Maximiser la marge O est équivalent à maximiser la somme des distances des deux classes par rapport à l'hyperplan. Ainsi la marge a l'expression mathématique suivante :

$$\begin{aligned} M &= \min_{x_i|y_i=1} \frac{w.x + b}{\|w\|} - \max_{x_i|y_i=-1} \frac{w.x + b}{\|w\|} \\ &= \frac{2}{\|w\|} \end{aligned} \quad (5.22)$$

Trouver HO revient donc à maximiser $\frac{2}{\|w\|}$. Ce qui est équivalent à minimiser $\frac{\|w\|^2}{2}$ sous la contrainte (5.21). Ceci est un problème de minimisation d'une fonction objectif quadratique avec contraintes linéaires. Ce problème peut se résoudre en utilisant les principes de Fermat, de Lagrange ou de Kuhn-Tucker [Cia94], [Fle87]. Le résumé de ces principes se trouve en annexe E.

En appliquant le principe de Kuhn-Tucker, on est amené à rechercher un point (w_0, b_0, α) . Le lagrangien correspondant à notre problème est :

$$L(w, b, \alpha) = \frac{1}{2} w \cdot w - \sum_{i=1}^m \alpha_i [y_i [(x_i \cdot w) + b] - 1] \quad (5.23)$$

Le lagrangien doit être minimal par rapport à w et b et maximal par rapport à $\alpha \geq 0$.

- $L(w, b, \alpha)$ est minimal par rapport à b :

$$\frac{\partial L(w, b, \alpha)}{\partial b} = 0 \Leftrightarrow \sum_{i=1}^m \alpha_i y_i = 0 \quad (5.24)$$

- $L(w, b, \alpha)$ est minimal par rapport à w :

$$\frac{\partial L(w, b, \alpha)}{\partial w} = 0 \Leftrightarrow w - \sum_{i=1}^m \alpha_i x_i y_i = 0 \quad (5.25)$$

- $L(w, b, \alpha)$ est maximal par rapport à $\alpha \geq 0$:

En remplaçant (5.24) et (5.25) dans le lagrangien (5.23) on aura :

$$L(w, b, \alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (5.26)$$

Ainsi le problème est de maximiser $L(w, b, \alpha)$ sous la contrainte :

$$\sum_{i=1}^m \alpha_i y_i = 0; \quad \alpha_i \geq 0 \quad (5.27)$$

Soit la solution $\alpha^0 = (\alpha_1^0, \dots, \alpha_m^0)$. D'après le théorème de Kuhn-Tucker une condition nécessaire et suffisante pour que α^0 soit optimal est :

$$\alpha_i^0 [y_i [(w_0 \cdot x) + b_0] - 1] = 0 \quad \text{pour } i = 1, \dots, m \quad (5.28)$$

Ce qui veut dire que : $\alpha_i^0 = 0$ où $[y_i [(w_0 \cdot x) + b_0] - 1] = 1$

Les Vecteurs Supports SV sont définis comme étant tout vecteur x_i tel que $y_i[(w_0 \cdot x_i) + b_0] = 1$. Ce qui est équivalent à :

$$VS = \{x_i / \alpha_i > 0\} \text{ pour } i = 1, \dots, m \quad (5.29)$$

Ainsi, w_0 et b_0 sont facilement calculés :

$$w_0 = \sum_{VS} \alpha_i^0 y_i x_i \quad (5.30)$$

$$b_0 = -\frac{1}{2}[(w_0 \cdot x^*(1)) + [(w_0 \cdot x^*(-1))]] \quad (5.31)$$

La fonction de classement $class(x)$ est défini par :

$$\begin{aligned} class(x) &= sign[(w_0 \cdot x) + b_0] \\ &= sign\left[\sum_{x_i \in VS} \alpha_i^0 y_i (x_i \cdot x) + b_0\right] \end{aligned} \quad (5.32)$$

Si $class(x)$ est inférieur à 0, x est de la classe -1 sinon il est de la classe 1.

b. Cas des données non linéairement séparables

Dans ce cas où les données sont non linéairement séparables, l'hyperplan optimal est celui qui satisfait les conditions suivantes:

- la distance maximale entre les vecteurs bien classés et l'hyperplan optimal doit être maximale.
- La distance entre les vecteurs mal classés et l'hyperplan optimal doit être minimale.

Pour formaliser ces conditions, des variables de pénalité non-négatives (ξ , $i = 1, \dots, m$) appelées encore variables d'écart sont introduites. Ces variables transforment l'inégalité (5.21) comme suit :

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i \text{ pour } i = 1, \dots, m \quad (5.33)$$

L'objectif est de minimiser la fonction suivante :

$$\Psi(w, E) = \frac{1}{2} w \cdot w + C \sum_{i=1}^m \xi_i \quad (5.34)$$

où C est un paramètre de régularisation. Il permet de définir l'importance aux erreurs. Cela mène à un problème légèrement différent de celui du cas des données linéairement séparables, à savoir maximiser le lagrangien donné par l'équation (5.26) par rapport à α_i sous les conditions suivantes :

$$\sum_{i=1}^m \alpha_i y_i = 0 \quad \text{avec} \quad 0 \leq \alpha_i \leq C \quad \text{pour } i = 1, \dots, m \quad (5.35)$$

Le calcul de la normale w_0 , du biais b_0 et de la fonction de classification $\text{class}(x)$ reste exactement le même que pour les données linéairement séparables.

c. Cas d'utilisation d'un classifieur non linéaire

Les classifieurs SVM à noyau utilisent le même principe de l'hyperplan optimal pour calculer une frontière entre des nuages de points. Cependant l'idée principale est cette fois de construire une surface de séparation linéaire dans un nouvel espace, l'espace des caractéristiques, qui correspond à une surface non- linéaire dans l'espace d'entrée.

Soit l'ensemble D' l'image de l'ensemble D , défini par la transformation π .

$$D' = \left\{ (\pi(x_i), y_i) \in \mathfrak{R}^p \times \{-1, 1\} \text{ pour } i = 1, \dots, m \mid p \geq n \right\} \quad (5.36)$$

En construisant un hyperplan optimal dans l'espace des caractéristiques suivant la technique expliquée dans le paragraphe précédent, la fonction de classement sera :

$$\text{class}(x) = \text{sign} \left[\sum_{x_i \in \mathcal{V}S} \alpha_i^0 y_i (\pi(x_i) \cdot \pi(x)) + b_0 \right] \quad (5.37)$$

Pour que le coût de calcul reste pratiquement inchangé et que le nombre de paramètres libres du système n'augmente pas, il faut que la fonction π satisfasse la condition suivante :

$$\pi(u).\pi(v) = K(u, v) \quad (5.38)$$

C'est à dire que le produit scalaire dans l'espace des caractéristiques va être représenté comme un noyau de l'espace d'entrée.

Plusieurs noyaux ont été utilisés par les chercheurs, en voici quelques uns :

- Le noyau linéaire

$$K(u, v) = u.v \quad (5.39)$$

- Le noyau polynomial

$$K(u, v) = [(u.v) + 1]^d \quad (5.40)$$

où d est le degré du polynôme à déterminer par l'utilisateur.

- Le noyau RBF (Radial Basis Function)

$$K(u, v) = \exp\left(-\frac{\|u - v\|^2}{2\sigma^2}\right) \quad (5.41)$$

où σ est à déterminer.

Ayant défini le noyau, la fonction de classement (5. 37) devient :

$$class(x) = sign\left[\sum_{x_i \in \mathcal{I}^S} \alpha_i^0 y_i K(x_i, x) + b_0\right] \quad (5.42)$$

5.5 Application aux événements de l'EMG utérin

Dans le cas de classification des événements de l'EMG utérin, cinq classes (contractions CT, ondes d'Alvarez (Alv), Mouvements actifs fœtaux (MAF), contractions LDBF, et bruit) ont été générées. Elles constituent les événements de la base REEL1 (annexe A). Chacune des classes est formée de 50 événements, 250 événements constituent l'ensemble d'apprentissage et les 250 autres événements forment l'ensemble test. Chaque événement est représenté par 30 paquets d'ondelettes. Chaque paquet est caractérisé par la variance de ses coefficients.

Comme expliqué au début de ce chapitre, nous disposons jusqu'à maintenant de quatre classes dites "physiologiques" dans l'EMG utérin : contractions, mouvements actifs fœtaux, ondes d'Alvarez, ondes LDBF. L'activité de fond constitue la cinquième classe. La classification

d'un groupe ou d'une classe est fondée sur l'estimation de sa probabilité d'appartenance à l'une ou l'autre des classes "physiologiques". La validation globale des algorithmes de classification a été menée en utilisant l'ensemble des données disponibles pour chaque classe. Les méthodes de classification : K Plus Proches Voisins, distance de Mahalanobis, réseaux de neurones et SVM ont été utilisées.

Les performances ont d'abord été calculées en utilisant uniquement la variance des coefficients des paquets choisis. Pour améliorer les performances, une caractéristique supplémentaire, la durée des événements, a été ajoutée. L'ensemble G1REEL1 (de la base REEL1 évoquée à l'annexe A) composée de 50 événements de chaque type a été utilisée pour l'ensemble d'apprentissage et les événements de l'ensemble G2REEL1 (de la base REEL1 évoquée à l'annexe A) sont utilisés pour l'ensemble test.

Chacun des événements est décomposé en paquets d'ondelettes jusqu'au niveau 4. Ce choix du niveau est justifié par le fait que le niveau 4 contient encore des paquets à bandes fréquentielles correspondant bien à certains événements de l'EMG utérin comme les mouvements actifs fœtaux et les ondes LDBF. Un tableau de 100×30 (100 événements et 30 valeurs de variance) est calculé pour chaque type d'événement. Comme c'est l'ondelette Symlet 5 qui a été choisie dans le cas de la détection, nous continuons à utiliser la même ondelette dans le cas de la classification, sachant qu'une étude reste à faire pour déterminer s'il existe une ondelette plus performante et même adaptée au problème de classification.

5.5.1 Sélection des paquets discriminants

Revenons au critère de discrimination défini à l'équation 5.10 :

$$R^n = \frac{\sum_w^{\wedge n}}{\sum_w^{\wedge n} + \sum_B^{\wedge n}}$$

où n est le numéro de paquet. $n = 1 \dots N$ avec N le nombre total de paquets.

Dans notre cas, $N = 30$. Les différentes valeurs de R^n ont été calculées. La figure 5.2 montre les valeurs de R^n par ordre croissant.

Les numéros des paquets par ordre croissant sont alors : 15, 7, 3, 1, 16, 8, 18, 17, 4, 10, 22, 21, 20, 23, 11, 24, 25, 5, 9, 29, 28, 12, 13, 14, 19, 2, 27, 26, 6, et 30. La sélection des paquets discriminants se fait en fixant par exemple un seuil pour déterminer les paquets ayant les plus

petites valeurs de R^n . Ce seuil sur les valeurs de R^n est explicite sur la figure 5.2. Un seuil égal à 0.7 a été choisi. Par conséquent les paquets 15, 7, 3, 1 et 16 ont été choisis comme étant les paquets ayant les caractéristiques les plus discriminantes. Pour démontrer la performance associée au choix de ce seuil, les méthodes de K Plus Proche Voisin et de distance de Mahalanobis ont été utilisées. Les pourcentages de bonne classification ont été calculés pour l'ensemble test. Pour ce test, chaque paquet n est caractérisé que par ses coefficients.

L'analyse prend d'abord le paquet le plus discriminant (15), puis les deux paquets les plus discriminants (15, 7) et ainsi de suite, jusqu'à prendre en compte l'ensemble des trente paquets. Le taux de bonne classification est calculé dans chaque cas et par les deux approches (Mahalanobis et KPPV).

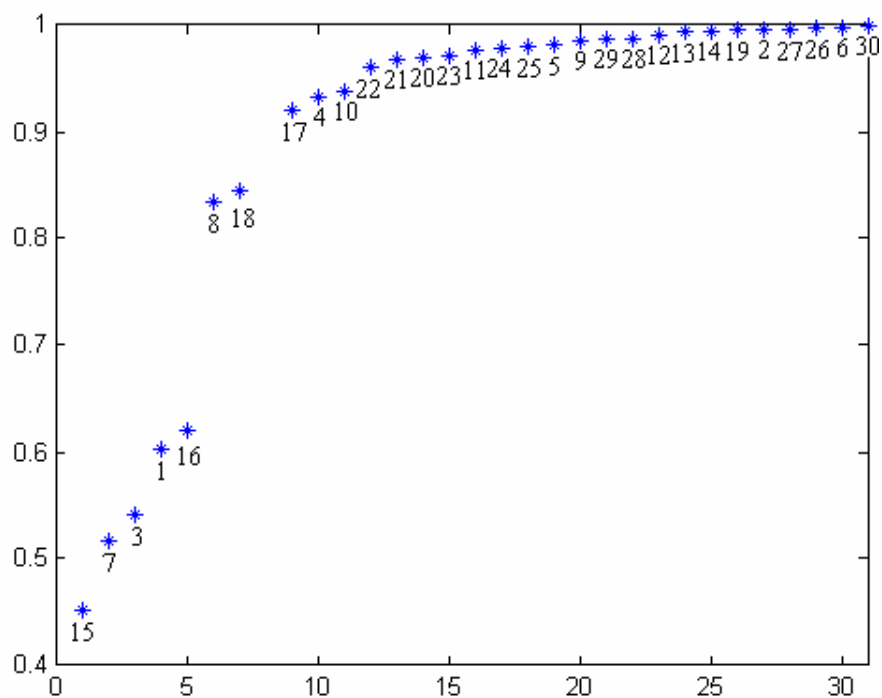


Figure 5.2 Ordre croissant des valeurs de R^n pour chaque paquet de numéro n . Axe des abscisses : unités arbitraires. Axe des ordonnées : valeurs de R^n .

5.5.2 Vérification de la performance du choix des paquets : utilisation de la distance de Mahalanobis

Le tableau 5.1 montre les pourcentages de bonne classification obtenus pour chaque classe d'événements. Il présente les numéros des paquets avec les taux de bonne classification en commençant par le paquet ayant la plus petite valeur de R^n qui est le paquet numéro 15, puis en ajoutant les paquets jusqu'au dernier (30).

Numéro des paquets	CT	Alv	MAF	LDBF	Bruit
15	0.5000	0.5400	0.6000	0.7000	0.6000
7	0.5000	0.5600	0.6000	0.7400	0.7400
3	0.5200	0.5400	0.6400	0.7400	0.8400
1	0.5200	0.5400	0.7000	0.6800	0.8600
16	0.5200	0.7000	0.7000	0.7600	0.9000
8	0.5000	0.7000	0.7000	0.7000	0.9000
18	0.4800	0.7000	0.7000	0.7000	0.9000
17	0.4400	0.7000	0.6800	0.6600	0.9600
4	0.4800	0.6800	0.6800	0.6200	0.9600
10	0.3800	0.6400	0.6800	0.5800	0.9600
22	0.3800	0.6600	0.6600	0.5400	0.9800
21	0.3200	0.5600	0.6000	0.5200	0.9800
20	0.2800	0.5400	0.6200	0.4400	0.9800
23	0.2200	0.5200	0.5800	0.4400	0.9800
11	0.2200	0.5000	0.5600	0.4400	0.9800
24	0.2200	0.5200	0.5200	0.3800	0.9800
25	0.2200	0.4800	0.5400	0.3400	0.9800
5	0.2200	0.4800	0.5400	0.3400	0.9800
9	0.2000	0.5000	0.5200	0.3200	0.9800
29	0.1800	0.4800	0.5200	0.3200	0.9800
28	0.1800	0.4800	0.5000	0.3200	0.9800
12	0.1800	0.4800	0.4600	0.2800	1.0000
13	0.1400	0.4600	0.4400	0.2800	1.0000
14	0.1600	0.4000	0.4400	0.2400	0.9800
19	0.1600	0.3600	0.4600	0.2000	0.9800
2	0.1400	0.4000	0.4800	0.1800	1.0000
27	0.1200	0.3800	0.4600	0.1400	1.0000
26	0.1200	0.4000	0.4400	0.1400	1.0000
6	0.1000	0.4200	0.4000	0.1400	1.0000
30	0.1000	0.4000	0.4000	0.1000	1.0000

Tableau 5.1. Pourcentages de bonne classification pour chaque type d'événements en commençant par le paquet à plus petite valeur de R^n jusqu'au paquet à plus grande valeur.

Les taux de bonne classification des paquets déjà identifiés comme étant les plus discriminants (15, 7, 3, 1 et 16) donnent à peu près les taux les plus élevés pour tous les événements. Le tableau 5.2 montre les taux de bonne classification en appliquant la méthode de distance de Mahalanobis uniquement sur les paquets 15 et 16, qui ont été obtenus après application du choix de la meilleure base.

Paquets	CT	Alv	MAF	LDBF	Bruit
15, 7, 3, 1,16	0.52	0.72	0.70	0.76	0.90
15, 16	0.66	0.50	0.74	0.82	0.70

Tableau 5.2. Pourcentages de bonne classification après réduction du nombre des paquets discriminants.

D'après le tableau 5.2, les résultats obtenus après réduction montre que les taux de bonne classification des ondes LDBF, mouvements actifs foetaux et contractions ont été augmentés, alors qu'ils sont moins bons pour les ondes d'Alvarez et le bruit. Ceci s'explique par le fait que les paquets 15 et 16 ayant les bandes fréquentielles [0-0.5] et [0.5-1] Hz pour une fréquence d'échantillonnage $F_e = 16$ Hz correspondent aux caractéristiques fréquentielles des ondes LDBF, MAF et contractions. Dans le cas des ondes Alvarez et du bruit qui correspondent aux hautes fréquences, ils ont perdu une partie de l'information qui est contenue dans les paquets 1, 3 et 7 et qui correspondent aux bandes fréquentielles [0-4],[0-2] et [0-1] Hz.

Comme les bandes de fréquence de tous les événements (sauf le bruit) sont incluses dans la bande [0-1] Hz (surtout les ondes Alvarez qui sont à des fréquences > 0.5 Hz), introduisons alors le paquet 7 (cette introduction constitue l'étape supplémentaire de l'algorithme du choix de la meilleure base pour la classification au paragraphe .4.2), et calculons de nouveau les taux de bonne classification. Le tableau 5.3 montre les résultats obtenus pour les paquets 15, 16 et 7.

Paquets	CT	Alv	MAF	LDBF	Bruit
15, 7, 3, 1,16	0.52	0.72	0.70	0.76	0.9
15, 16	0.66	0.50	0.74	0.82	0.7
15, 16, 7	0.68	0.70	0.78	0.82	0.84

Tableau 5.3. Pourcentages de bonne classification après le choix final des paquets discriminants.

En comparant les valeurs de deux tableaux 5.2 et 5.3, on peut conclure que les paquets les plus discriminants pour la méthode de distance de Mahalanobis sont 7, 15 et 16. Le paragraphe suivant va permettre de comparer ces résultats à une méthode plus locale, celle des K-PPV.

5.5.3 Vérification de la performance du choix des paquets : utilisation de la méthode des K plus Proches Voisins

Le tableau 5.4 présente les pourcentages de bonne classification de chacun des événements en commençant par le paquet le plus discriminant (15), puis les 2 paquets les plus discriminants (15, 7) et ainsi de suite jusqu'à prendre en compte l'ensemble des trente paquets (comme cela a déjà été fait dans le cas de la distance de Mahalanobis). Dans ce cas de classification, le choix de la valeur de K n'est pas important car le but est de vérifier si les paquets choisis (15, 7, 3, 1 et 16) contribuent à discriminer entre les différentes classes. La valeur absolue de la performance n'a donc pas ici d'importance primordiale. Cependant, pour chaque ensemble test, les taux de bonne classification ont été calculés pour les valeurs de K allant de 1 jusqu'à 50 et pour tous les paquets obtenus en ajoutant un paquet selon l'ordre obtenu à la figure 5.2. La valeur de $K = 14$ a donné les meilleurs pourcentages de bonne classification pratiquement pour tous les événements et surtout pour les contractions.

Les pourcentages de bonne classification présentés au tableau 5.4 sont alors calculés pour $K = 14$. En examinant le tableau 5.4, nous constatons que les pourcentages de bonne classification sont pratiquement les meilleurs pour les paquets 15, 7, 3, 1 et 16. Le tableau 5.5 présente les taux de bonne classification en utilisant les paquets obtenus par sélection de la meilleure base (paquets 15 et 16) et après correction (paquets 15, 16 et 7). Nous remarquons que les résultats sont semblables dans les deux cas : paquets (15, 7, 3, 1 et 16) et paquets (15, 16 et 7).

5.5.4 Conclusion

De ce qui précède, nous pouvons conclure que les paquets (15, 7, 3, 1 et 16) sont les plus discriminants. Les paquets retenus après réduction (15, 16 et 7) ont conservé le pouvoir discriminant et ont donné des pourcentages de bonne classification améliorés pour certains événements. On peut alors retenir les paquets numéros (15, 16 et 7) et les utiliser comme des entrées pour la classification finale. Dans ce qui suit, la durée sera utilisée comme une caractéristique supplémentaire pour améliorer les taux de bonne classification.

Numéro des paquets	CT	Alv	MAF	LDBF	Bruit
15,7	0.64	0.36	0.66	0.88	0.34
3	0.62	0.4	0.68	0.88	0.46
1	0.62	0.4	0.68	0.88	0.46
16	0.72	0.48	0.68	0.9	0.48
8	0.7	0.48	0.68	0.9	0.48
18	0.7	0.5	0.68	0.9	0.48
17	0.7	0.48	0.68	0.9	0.48
4	0.7	0.48	0.68	0.9	0.48
10	0.7	0.48	0.68	0.9	0.48
22	0.7	0.48	0.68	0.9	0.48
21	0.7	0.48	0.68	0.9	0.48
20	0.72	0.48	0.68	0.9	0.48
23	0.72	0.48	0.68	0.9	0.48
11	0.72	0.48	0.68	0.9	0.48
24	0.72	0.48	0.68	0.9	0.48
25	0.72	0.5	0.68	0.9	0.48
5	0.72	0.5	0.68	0.9	0.5
9	0.72	0.5	0.68	0.9	0.5
29	0.72	0.5	0.68	0.9	0.52
28	0.72	0.5	0.68	0.9	0.54
12	0.72	0.5	0.68	0.9	0.54
13	0.72	0.5	0.68	0.9	0.54
14	0.72	0.5	0.68	0.9	0.54
19	0.72	0.5	0.68	0.9	0.54
2	0.72	0.5	0.68	0.9	0.54
27	0.72	0.5	0.68	0.9	0.54
26	0.72	0.5	0.68	0.9	0.56
6	0.72	0.5	0.68	0.9	0.56
30	0.72	0.5	0.68	0.9	0.56

Tableau 5.4. Pourcentages de bonne classification en commençant par le paquet à plus petite valeur de R^n jusqu'au paquet à plus grande valeur. Méthode de KPPV.

5.6 Classification à partir des paquets sélectionnés

Les paquets qui seront utilisés sont donc ici les paquets 15, 16 et 7. Les pourcentages de bonne classification sont calculés pour les quatre méthodes de classification présentées dans

les paragraphes précédents, mais en introduisant cette fois la durée des événements pour améliorer la performance.

Paquets	CT	Alv	MAF	LDBF	Bruit
15, 7, 3, 1,16	0.72	0.48	0.68	0.9	0.48
15, 16	0.72	0.4	0.68	0.88	0.40
15, 16, 7	0.72	0.46	0.76	0.9	0.48

Tableau 5.5. Pourcentages de bonne classification pour les différentes approches de réduction.

5.6.1 Distance de Mahalanobis et des K Plus Proches Voisins

Le tableau 5.6 présente les taux de bonne classification en utilisant les paquets (15, 16 et 7) et en introduisant la durée des événements comme une caractéristique supplémentaire.

D'après le tableau 5.6, on remarque que tous les pourcentages de bonne classification ont été améliorés, ce qui confirme que l'introduction de la durée des événements est un bon choix pour améliorer la discrimination entre les événements de l'EMG utérin.

Méthode	CT	Alv	MAF	LDBF	Bruit
Distance de Mahalanobis	0.84	0.9	0.82	1	0.94
KPPV	0.80	0.98	0.96	1	0.74

Tableau 5.6. Pourcentages de bonne classification en introduisant la durée des événements pour les méthodes KPPV et distance de Mahalanobis.

5.6.2 Réseaux de neurones

Le réseau de neurone utilisé est le Perceptron Multi-Couche à retro-propagation (MLP). Pour spécifier la structure du réseau, il est nécessaire de choisir le nombre de couches et le nombre de neurones sur chaque couche. Le nombre d'entrées du réseau, de même que le nombre de neurones sur la couche de sortie est fixé par les spécifications du problème que l'on veut résoudre avec ce réseau. Rappelons que 5 classes ont été appliquées à l'entrée du réseau de neurones, avec les caractéristiques obtenues à partir des paquets 15, 16 et 7.

Le perceptron multi-couche utilisé dans notre cas, est formé d'une couche cachée formée de 5 neurones et d'une couche de sortie (où sont présentées les sorties calculées par le MLP).

La fonction d'activation utilisée pour la couche cachée est la fonction tangente hyperbolique :

$$f(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}}$$

La fonction d'activation utilisée pour la couche de sortie est la fonction linéaire.

$$f(n) = n$$

Le tableau 5.7 montre que de bons résultats ont été obtenus. L'introduction de la durée a largement amélioré les pourcentages de bonne classification surtout pour les ondes d'Alvarez et les ondes MAF.

	CT	Alv	MAF	LDBF	Bruit
Sans Durée	0.78	0.72	0.78	0.94	0.86
Avec durée	0.86	0.96	0.98	1.00	0.86

Tableau 5.7. Pourcentages de bonne classification en utilisant les paquets 15, 16 et 7 sans et avec la durée des événements. Le réseau de neurones MLP est utilisé.

5.6.3. Machines à Vecteurs Supports

Pour discriminer entre les différents événements utérins, la méthode de SVM multiclasse a été utilisée. La méthode adoptée dans notre cas est celle évoquée en [Dua05] et appelée WTA-SVM. Etant donné M classes, cette méthode construit M classifieurs. La $i^{\text{ème}}$ sortie est effectuée en prenant les éléments d'une classe donnée comme ayant les étiquettes positives (+1) et les éléments de toutes les autres classes à étiquettes négatives (-1). Pour M classes, il faut définir M classifieurs SVM.

Dans notre application, le noyau choisi est RBF (Radial Basis Function).

Deux paramètres sont à régler : σ et C (paramètre de régularisation). L'annexe F contient les tableaux des pourcentages de bonne classification de tous les événements en prenant comme valeurs de σ (0.1, 0.2, 0.4, 0.6, 0.8, 1, 5 et 10) et comme valeurs de C (0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000 et Inf). Les pourcentages de bonne classification sont calculés en utilisant les caractéristiques des paquets 15, 16 et 7 d'une part et en ajoutant la durée des événements d'autre part. Le tableau 5.8 est tiré de l'annexe F. Il présente les valeurs σ et C qui donnent les meilleurs taux de bonne classification pour tous les événements, dans les deux cas avec durée et sans durée.

Le tableau 5.8 montre que la méthode SVM a donné de bons résultats dans les deux cas (avec et sans durée). Pour les ondes d'Alvarez, le taux de bonne classification était 0.64 sans la durée, et il a été bien amélioré (0.86) après l'introduction de la durée.

	CT	Alv	MAF	LDBF	Bruit
Sans Durée	$\sigma=5, C=Inf$ 0.82	$\sigma=10, C=Inf$ 0.64	$\sigma=5, C=Inf$ 0.80	$\sigma=1, C=Inf$ 0.88	$\sigma=5, C=100$ 0.82
Avec Durée	$\sigma=5, C=Inf$ 0.88	$\sigma=5, C=1$ 0.86	$\sigma=5, C=1$ 0.96	$\sigma=10, C=1$ 0.98	$\sigma=0.6, C=1$ 0.90

Tableau 5.8. Pourcentages de bonne classification dans le cas SVM, avec le choix optimal de σ et C .

5.7 Conclusion

L'obtention des paquets les plus discriminants à partir de l'arbre obtenu par décomposition en paquets d'ondelettes, est guidée par un critère répondant à l'objectif de classification. Le rapport entre la variance intra classe et la variance totale paraît un bon critère de choix des meilleurs paquets discriminants.

Le choix des paquets à mêmes bandes fréquentielles que les événements de l'EMG utérin améliore la performance de cette méthode dans le choix des paquets significatifs en terme de classification d'événements. L'identification "physiologique" des événements obtenus précédemment a été réalisée tout d'abord par deux méthodes de classification (K Plus Proches Voisins et distance de Mahalanobis) pour montrer l'efficacité des paquets discriminants à différencier entre les différents événements de l'EMG utérin, et le bruit. L'intérêt de la réduction des paquets discriminants a été vérifiée par ces deux méthodes. L'approche de réduction utilisée dans le cas de la détection a été légèrement modifiée dans le cas de la classification. Les paquets qui sont à bandes fréquentielles très proches de celles des événements des classes ont été conservés afin de permettre une discrimination plus efficace. Ça a été le cas du paquet 7 car très proche du contenu fréquentiel des ondes d'Alvarez.

Cette procédure heuristique souligne néanmoins les limites de la méthode de sélection de la meilleure base. La réduction de redondance risque en effet parfois d'entraîner une dégradation notable des performances de classification vis-à-vis d'un groupe spécifique d'événements.

Après le choix des meilleurs paquets discriminants, une caractéristique supplémentaire correspondant à la durée des événements a été ajoutée aux entrées des classifieurs pour améliorer les performances de bonne classification.

La valeur du K dans la méthode des K plus Proches Voisins a été choisie comme étant la valeur qui donne le taux de bonne classification le plus élevé, en particulier pour les contractions.

Le réseau de neurone utilisé est le Perceptron Multi-Couche à retro-propagation (MLP). Le principal problème lié à son utilisation est le choix des fonctions d'activation des couches cachée et de sortie, en plus de la définition du nombre de neurones cachés. Dans notre application la fonction tangente hyperbolique a été utilisée pour la couche cachée et la fonction linéaire a été utilisée pour la couche de sortie. De bons résultats ont été obtenus avec ou sans utilisation de la durée.

L'approche multiclasse de la méthode SVM utilisée est celle qui construit un modèle SVM pour chaque groupe d'événement permettant de le discriminer par rapport à l'ensemble des autres groupes.

Les paramètres choisis pour chaque système sont ceux qui donnent les meilleurs pourcentages de bonne classification. Ces paramètres varient selon que la durée des événements est utilisée ou non.

Les pourcentages de bonne classification ont été bien améliorés par l'introduction de la durée, quelle que soit la méthode de classification.

Pour les 4 méthodes utilisées, de bons résultats ont été obtenus. En moyenne plus de 85% des événements ont été bien classifiés, quel que soit le terme de grossesse.

Conclusion générale

Nous avons introduit ce travail en présentant deux approches classiques de décomposition du signal, la décomposition discrète en ondelettes et la décomposition en paquets d'ondelettes fondées sur la construction de sous-bases de fonctions organisées en paquets.

Nous avons défini un algorithme de recherche d'une meilleure base de paquets d'ondelettes pour la détection et, par suite, pour la classification. L'efficacité de l'algorithme dépend d'un paramètre de seuillage des valeurs du critère de détection ou de classification. Des tests réalisés sur des signaux simulés et réels ont montré de bonnes performances.

Le chapitre 1 a été consacré à une revue des données physiologiques nécessaires d'une part à la bonne compréhension des événements en jeu, d'autre part à une définition claire du problème posé. C'est ainsi que nous avons rappelé que les objectifs généraux de l'étude consistent à prévenir les accouchements prématurés. Nous avons décrit l'EMG utérin comme une intégration filtrée des activités électriques émises par des cellules utérines sous-jacentes.

L'utilisation de méthodes adaptatives pour le traitement du signal EMG a été justifiée. L'approche du traitement du signal et de la reconnaissance de formes s'est appuyée sur une base de données disponibles des événements et des signaux femmes, en utilisant une modélisation spécifique faisant apparaître les caractéristiques qui permettent de détecter et classifier les événements contenus dans le signal. Cette modélisation a constitué la première étape de l'étude.

Dans le chapitre 2, deux types de modélisations ont été présentés : la première est la décomposition discrète en ondelettes, la deuxième la décomposition en paquets d'ondelettes. L'analyse multi résolution et le calcul des coefficients dans le cas de DWT et WPT, ainsi que les différents algorithmes du choix de la meilleure base associés aux divers critères utilisés et répondant aux objectifs de la sélection des meilleurs paquets ont été présentés dans ce chapitre.

Après la modélisation du signal en paquets d'ondelettes, nous avons proposé dans le chapitre 3 un algorithme du choix de la meilleure base consistant à identifier les paquets révélant au mieux l'existence d'événements dans une classe donnée de signaux.

Comme la distance de Kullback Leibler est un bon outil statistique pour décider de la capacité d'un paquet à mettre en évidence les différences de caractéristiques statistiques entre les catégories d'événements, elle a été utilisée comme un critère du choix de la meilleure base pour la détection.

Les paramètres de la loi gaussienne généralisée pour les événements et les signaux EMG utérins ont été estimés. Le calcul de la statistique de Kolmogorov Smirnov D_{max} a montré que les événements et les signaux EMG ainsi que les coefficients des paquets d'ondelettes associés suivent une loi gaussienne généralisée alors que le bruit suit la loi gaussienne standard.

Comme l'idée de base de ce chapitre est de choisir les meilleurs paquets pertinents en détection, la distance de K-L est estimée pour chaque paquet et la statistique de K-S est calculée pour montrer l'écart entre histogramme relatif aux distances de K-L estimées et la loi exponentielle ayant la même espérance que celle de la loi de la distance de K-L en l'absence d'événements. Les valeurs de D_{max} sont calculées pour chaque paquet et triées en ordre décroissant, les paquets ayant la valeur de D_{max} plus grande qu'un certain seuil ont été choisis comme étant les paquets qui révèlent la présence d'événements. Le seuil sur les valeurs D_{max} était très clair, ce qui a permis de choisir simplement les meilleurs paquets. La méthodologie proposée a été appliquée aux signaux simulés et aux signaux réels. Pour les signaux EMG, 4 paquets ont été choisis, correspondant bien au contenu fréquentiel des événements de l'EMG. Comme l'arbre contient des paquets redondants, l'étape de réduction de la redondance a permis de diminuer le nombre de paquets sélectionnés sans modifier la performance globale. L'avis de l'expert sur la sélection des meilleurs paquets a confirmé le choix effectué grâce à l'application de la méthode de la sélection de la meilleure base.

Comme la loi exponentielle est une approche approximative de la distribution de la distance de K-L estimée \hat{K} , il pourrait être intéressant dans l'avenir de rechercher des modèles des lois plus proches de la distribution de \hat{K} que la loi exponentielle. De plus, il serait souhaitable d'élargir la base de données des événements et des signaux femmes.

Dans le chapitre 4, le principe retenu pour la détection est l'application d'algorithmes séquentiels sur les coefficients (et non sur les signaux reconstruits) des niveaux d'échelles

(dans le cas de décomposition discrète en ondelettes) ou sur les meilleurs paquets préalablement choisis (dans le cas de décomposition en paquets d'ondelettes). La détection des événements de l'EMG utérin est effectuée sur les coefficients des paquets sélectionnés. Les instants de changement ont été corrigés en utilisant des formules de correction déjà appropriées. Après l'obtention des instants de changement corrigés, une procédure de fusion a été appliquée pour redéfinir les vraies valeurs des instants de détection en référence au signal original. Les probabilités de bonne détection et de fausse alarme ont montré la bonne performance du choix de la meilleure base de détection et la pertinence de la correction et de la fusion effectuées sur les instants détectés. La détection en temps réel a donné de bons résultats également. Elle permet de détecter les événements sans attendre l'acquisition de tout le signal. La mesure de performance étant très liée aux avis des experts, il paraît important pour l'avenir de compléter le pool d'experts pour la localisation précise des événements de l'EMG utérin.

Le barycentre énergétique de la réponse impulsionnelle du filtre passe bas ou passe haut a été utilisé pour étudier l'influence de la répartition des coefficients en passant d'une échelle à l'autre. Malgré les bons résultats obtenus, il serait utile d'essayer avec la valeur maximale (ou autres valeurs) de la réponse impulsionnelle du filtre passe bas ou passe haut au lieu du barycentre énergétique. D'après les études effectuées, l'ondelette Symlet 5 a été retenue (parmi les ondelettes orthogonales).

L'étude du seuil par le tracé des courbes COR, pour décider de la valeur optimale pour les signaux femmes, est un travail qui mériterait d'être effectué également. Cependant, l'objectif de notre étude n'était pas l'algorithme de détection lui-même, mais le support sur lequel il est appliqué.

De même il resterait à appliquer d'autres algorithmes séquentiels de détection (autres que DCS) et à retenir des algorithmes aussi simples que possible surtout dans le cas de détection en temps réel.

La sélection de la meilleure base pour la classification est abordée au chapitre 5. Le rapport entre la variance intra classe et la variance totale est utilisé comme un critère de discrimination. Il est appliqué sur l'ensemble d'apprentissage des événements EMG utérins.

Pour cette application, l'ondelette Symlet 5 a été conservée. D'autres études concernant le choix de l'ondelette seraient intéressantes à développer. La capacité des meilleurs paquets à classifier les événements a été montrée en utilisant les méthodes de K Plus Proches Voisins et la distance de Mahalanobis. D'autres méthodes comme les réseaux de neurones et la méthode

SVM ont été utilisées avec comme entrées les paramètres issus des meilleurs paquets discriminants. La durée des événements a été introduite pour améliorer les pourcentages de bonne classification.

De bons résultats ont été obtenus : en moyenne plus de 85% des événements sont bien classifiés quelque soit le terme de grossesse.

Comme cela a déjà été précisé dans le cinquième chapitre, le domaine de la classification est très vaste et le nombre de méthodes développées très important. Bien que nous ayons testé quatre méthodes, il serait intéressant de tester d'autres méthodes sur les jeux de données disponibles, afin de choisir la plus adaptée et d'approfondir la recherche entamée sur la notion de classification supervisée.

En plus des perspectives déjà citées, il serait intéressant d'explorer d'autres types d'ondelettes aux caractéristiques fréquentielles plus sélectives, telles que les ondelettes splines cubiques orthogonales, de manière à faciliter la sélection des paquets d'une part, et à améliorer la qualité de détection d'autre part, ceci ayant pour conséquence de fournir un détecteur plus efficace. En d'autres termes, créer une ondelette adaptée aux événements de l'EMG et au décalage des coefficients serait important.

Ce travail nous laisse penser que le traitement et l'analyse des signaux EMG utérins peuvent être un outil d'investigation de la prévention de l'accouchement prématuré, voire de l'aide au diagnostic. Cependant, il est indispensable que tout le travail soit validé par une interprétation médicale qui puisse juger de la pertinence de l'étape finale de l'analyse. Enfin, seule l'application dans des situations réelles permettra d'apprécier l'utilité médicale de ces résultats.

Annexe A

Bases des signaux utilisées

Pour la vérification des méthodes développées dans ce travail, plusieurs bases des signaux ont été utilisées. Certaines bases contiennent des signaux synthétiques, d'autres contiennent des signaux réels de l'EMG utérin. Définissons alors les bases suivantes :

REEL1

Cette base est constituée de deux ensembles d'apprentissage et de test. Le premier ensemble s'appelle G1REEL1, alors que le deuxième s'appelle G2REEL1.

Chaque ensemble, formé des événements décomposés, comprend :

- 50 événements de contraction.
- 50 événements de mouvements actifs fœtaux.
- 50 événements d'ondes d'Alvarez.
- 50 événements d'ondes LDBF (Longue Durée Basse Fréquence).
- 50 segments de bruit réel (ligne de base).

REEL2

Cette base est formée des enregistrements réels de longue durée. Elle comprend 20 signaux réels qui ont été enregistrés sur des femmes à plusieurs termes de grossesse et au cours de l'accouchement. Ces enregistrements ont été effectués en hôpital et supervisés par l'équipe du Pr. Catherine Marque de l'Université de Technologie de Compiègne.

Les signaux ont été recueillis à l'aide d'une paire d'électrodes bipolaires apposées sur l'axe vertical médian de la paroi abdominale, en regard de l'utérus. Ces électrodes de surface étaient de type Ag/AgCl d'un diamètre de 8 mm. Les signaux étaient ensuite filtrés entre 0.2 et 6 Hz par un filtre analogique de façon à éliminer la composante continue et les artefacts dus au secteur, puis amplifiés. Ils étaient ensuite numérisés avec une fréquence d'échantillonnage de 16 Hz.

SIM1

Cette base est constituée de deux ensembles. Le premier ensemble d'apprentissage s'appelle G1SIM1, alors que le deuxième s'appelle G2SIM1. Chaque ensemble est composé de 1000 signaux simulés.

Chaque signal est composé de 4 segments (2000 points) correspondant à différentes bandes fréquentielles dont les bornes sont définies par référence à la moitié de la fréquence d'échantillonnage $F_e / 2$ ([0.01 ; 0.4], [0.01 ; 0.6], [0.01 ; 0.8], et [0.01 ; 0.999] Hz). La longueur totale du signal en nombre de points est 8000.

SIM2

Cette base est constituée de signaux composés de 3 segments (1000 points) correspondant à différentes bandes fréquentielles dont les bornes sont définies par référence à la moitié de la fréquence d'échantillonnage $F_e / 2$ ([0.01 ; 0.25], [0.26 ; 0.5], [0.51 ; 0.75]). La longueur totale du signal en nombre de points est 3000.

SIM3

Cette base contient des signaux obtenus également par simulation. 1000 signaux ont été générés. Chaque signal est composé de deux segments de 2000 points de bruit blanc gaussien suivant les lois $N(0,1)$ et $N(0,4)$ respectivement. L'instant de changement entre les deux distributions est 2000 (numéro du point).

REELSIM

Cette base est constituée de deux ensembles. Le premier ensemble d'apprentissage s'appelle G1REELSIM, alors que le deuxième s'appelle G2REELSIM.

Chaque ensemble est composé de 100 signaux. Chaque signal comprend 10 événements de différents types appartenant à différentes femmes et termes (contractions, ondes d'Alvarez, ondes LDBF, mouvements actifs fœtaux...) identifiés par un expert, ces événements ont été chaînés, contaminés par du bruit blanc (SNR = 10 db) et séparés par du bruit blanc (SNR = 10db).

Annexe B

Estimation des paramètres d'échelle et de forme α et β pour une distribution gaussienne généralisée GGD

Plusieurs approches ont été adoptées pour estimer α et β [Sha95], [Do02], [Var89]. Récemment, *M.N. Do* et *M. Vetterli* ont proposé une estimation basée sur le principe du maximum de vraisemblance [Do02]. Considérons une observation $\mathbf{x} = (x_1, \dots, x_N)$ où les différents échantillons x_i sont considérés indépendants. Définissons la fonction de vraisemblance :

$$L(\mathbf{x}; \alpha, \beta) = \log \prod_{i=1}^N f(x_i; \alpha, \beta) \quad (\text{B.1})$$

où l'expression de f est donné par la Densité Gaussienne Généralisée GGD :

$$f(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta} \quad (\text{B.2})$$

Les paramètres α et β doivent être estimés. Le principe du maximum de vraisemblance nous conduit aux équations ci-dessous :

$$\left\{ \begin{array}{l} \frac{\partial L(\mathbf{x}; \alpha, \beta)}{\partial \alpha} = -\frac{L}{\alpha} + \sum_{i=1}^L \frac{\beta |x_i|^\beta \alpha^{-\beta}}{\alpha} = 0 \\ \frac{\partial L(\mathbf{x}; \alpha, \beta)}{\partial \beta} = \frac{L}{\beta} + \frac{L\Psi(1/\beta)}{\beta^2} - \sum_{i=1}^L \left(\frac{|x_i|}{\alpha} \right)^\beta \log \left(\frac{|x_i|}{\alpha} \right) = 0 \end{array} \right. \quad (\text{B.3})$$

où $\Psi(\cdot)$ est la fonction digamma [Abr70], i.e., $\Psi(z) = \Gamma'(z)/\Gamma(z)$.

Ce système d'équations admet une solution unique appelée l'estimé au sens du maximum de vraisemblance du couple de paramètres (α, β) . Quand on fixe le paramètre $\beta > 0$, l'estimé de α est alors donné par :

$$\hat{\alpha} = \frac{\beta}{L} \left(\sum_{i=1}^L |x_i|^\beta \right)^{1/\beta} \quad (\text{B.4})$$

L'estimé du paramètre de forme $\hat{\beta}$ est la solution de l'équation transcendante ci-dessous :

$$g(\hat{\beta}) = 1 + \frac{\Psi(1/\hat{\beta})}{\hat{\beta}} - \frac{\sum_{i=1}^L |x_i|^{\hat{\beta}} \log|x_i|}{\sum_{i=1}^L |x_i|^{\hat{\beta}}} + \frac{\log \left(\frac{\hat{\beta}}{L} \sum_{i=1}^L |x_i|^{\hat{\beta}} \right)}{\hat{\beta}} = 0 \quad (\text{B.5})$$

Cette équation est fortement et clairement non linéaire. Cependant, elle peut être résolue grâce à l'algorithme de Newton-Raphson [Kay93].

Il est important de noter que la méthode de Newton-Raphson est une procédure itérative qui essaie d'approximer la racine d'une fonction g donnée. Plus précisément, cette méthode dérive du fameux développement en séries de Taylor d'une fonction au voisinage d'un point :

$$g(x + \delta) = g(x) + \delta g'(x) + \frac{\delta^2}{2} g''(x) + \dots \quad (\text{B.6})$$

Pour les petites valeurs de δ et pour des fonctions g assez régulières, les termes d'ordre élevé dans le développement de Taylor peuvent être considérés comme négligeables. Ainsi, $g(x + \delta) = 0$ donne :

$$\delta \approx -\frac{g(x)}{g'(x)} \quad (\text{B.7})$$

Cette approximation nous amène à adopter l'algorithme de gradient pour trouver itérativement β_k :

$$\beta_{k+1} = \beta_k - \frac{g(\beta_k)}{g'(\beta_k)} \quad (\text{B.8})$$

où k est l'indice de l'itération.

Soit :

$$\begin{aligned}
 g'(\beta) = & -\frac{\Psi(1/\beta)}{\beta^2} - \frac{\Psi'(1/\beta)}{\beta^3} + \frac{1}{\beta^2} - \frac{\sum_{i=1}^N |x_i|^\beta (\log|x_i|)^2}{\sum_{i=1}^N |x_i|^\beta} + \frac{\left(\sum_{i=1}^N |x_i|^\beta \log|x_i|\right)^2}{\left(\sum_{i=1}^N |x_i|^\beta\right)^2} \\
 & + \frac{\sum_{i=1}^N |x_i|^\beta \log|x_i|}{\beta \sum_{i=1}^n |x_i|^\beta} - \frac{\log\left(\frac{\beta}{N} \sum_{i=1}^N |x_i|^\beta\right)}{\beta^2}
 \end{aligned} \tag{B.9}$$

où $\Psi'(z)$ est connu comme étant le premier polygamma ou la fonction trigamma.

La convergence est assurée à condition de prendre des conditions initiales correctes. Pour le cas de GGD, il est possible de démontrer que le rapport de la moyenne absolue et l'écart type est une fonction croissante de β :

$$F_M(\beta) = \frac{\Gamma(2\beta)}{\sqrt{\Gamma(1/\beta)\Gamma(3/\beta)}} \tag{B.10}$$

Si on considère que la moyenne et la variance sont estimées par :

$$m_1 = \frac{1}{N} \sum_{n=1}^N |x_n| \quad \text{et} \quad m_2 = \frac{1}{N} \sum_{i=1}^N x_i^2 \tag{B.11}$$

respectivement, alors β peut être estimé par résoudre l'équation suivante :

$$\bar{\beta} = F_M^{-1}\left(\frac{m_1}{\sqrt{m_2}}\right) \tag{B.12}$$

Dans l'implémentation pratique, la solution de l'équation (B.11) peut être rapidement déterminée par interpolation et un tableau "look-up table" dont les entrées sont les valeurs correspondantes de $\frac{m_1}{m_2}$ et $\bar{\beta}$. Finalement, la valeur initiale $\beta_0 = \bar{\beta}$ de l'estimateur de

maximum de vraisemblance $\hat{\beta}$ est affinée avec les premières étapes de l'algorithme de Newton-Raphson. Notons que 3 itérations suffisent pour converger. Le paramètre d'échelle α

est alors obtenu en réinjectant la valeur estimée de β dans l'équation donnant α (équation B.4). Ainsi, on a estimé à la fois les paramètres d'échelle et de forme.

Annexe C

La distance de K-L entre deux densités de probabilités f_{θ_1} et f_{θ_2} d'une variable aléatoire X , est définie par l'expression suivante :

$$D(f(X, \theta_1), (f(X, \theta_2))) = \int f(x, \theta_1) \log \frac{f(x, \theta_1)}{f(x, \theta_2)} dx. \quad (C.1)$$

L'expression de la densité gaussienne généralisée utilisée est :

$$f(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta} \quad (C.2)$$

où $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$ avec $z > 0$.

L'idée est de développer l'expression (C.1) en introduisant l'expression de la fonction de densité gaussienne généralisée (C.2).

Dans ce cas $D(f(X, \theta_1), (f(X, \theta_j)))$ s'écrit $D(f(;\alpha_1, \beta_1); f(;\alpha_2, \beta_2))$. Pour simplifier on utilisera la notion simplifiée D . En substituant (C.2) dans (C.1), l'équation (C.1) devient :

$$D = \int_{-\infty}^0 \frac{\beta_1}{2\alpha_1\Gamma(1/\beta_1)} e^{-(x/\alpha_1)^{\beta_1}} \log \left(\frac{\frac{\beta_1}{2\alpha_1\Gamma(1/\beta_1)} e^{-(x/\alpha_1)^{\beta_1}}}{\frac{\beta_2}{2\alpha_2\Gamma(1/\beta_2)} e^{-(x/\alpha_2)^{\beta_2}}} \right) dx + \int_0^\infty \frac{\beta_1}{2\alpha_1\Gamma(1/\beta_1)} e^{-(x/\alpha_1)^{\beta_1}} \log \left(\frac{\frac{\beta_1}{2\alpha_1\Gamma(1/\beta_1)} e^{-(x/\alpha_1)^{\beta_1}}}{\frac{\beta_2}{2\alpha_2\Gamma(1/\beta_2)} e^{-(x/\alpha_2)^{\beta_2}}} \right) dx \quad (C.3)$$

En substituant $x = -x$ dans la première partie de l'équation C.3, D s'écrit alors :

$$\begin{aligned}
D &= \int_0^{\infty} \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} e^{-(x/\alpha_1)^{\beta_1}} \log \left(\frac{\frac{\beta_1}{2\alpha_1 \Gamma(1/\beta_1)} e^{-(x/\alpha_1)^{\beta_1}}}{\frac{\beta_2}{2\alpha_2 \Gamma(1/\beta_2)} e^{-(x/\alpha_2)^{\beta_2}}} \right) dx \\
&= \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \left(\log \frac{\beta_1 \alpha_2 \Gamma(1/\beta_2)}{\beta_2 \alpha_1 \Gamma(1/\beta_1)} \right) \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} dx \\
&\quad - \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\frac{x}{\alpha_1} \right)^{\beta_1} dx \\
&\quad + \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\frac{x}{\alpha_2} \right)^{\beta_2} dx
\end{aligned} \tag{C.4}$$

Soit

$$D = P_1 + P_2 + P_3 \tag{C.5}$$

Rappelons que $\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt$ avec $z > 0$. Prenons $t = \left(\frac{x}{\alpha_1} \right)^{\beta_1}$, alors $dt = \frac{\beta_1}{(\alpha_1)^{\beta_1}} x^{(\beta_1-1)} dx$.

Calculons $\Gamma(1/\beta_1)$,

$$\begin{aligned}
\Gamma(1/\beta_1) &= \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\left(\frac{x}{\alpha_1} \right)^{\beta_1} \right)^{(1/\beta_1-1)} \frac{\beta_1}{(\alpha_1)^{\beta_1}} x^{(\beta_1-1)} dx \\
&= \frac{\beta_1}{\alpha_1} \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} dx
\end{aligned} \tag{C.6}$$

Par conséquent,

$$P_1 = \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \left(\log \frac{\beta_1 \alpha_2 \Gamma(1/\beta_2)}{\beta_2 \alpha_1 \Gamma(1/\beta_1)} \right) \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} dx = \log \left(\frac{\beta_1 \alpha_2 \Gamma(1/\beta_2)}{\beta_2 \alpha_1 \Gamma(1/\beta_1)} \right) \tag{C.7}$$

$$P_2 = \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\frac{x}{\alpha_1} \right)^{\beta_1} dx = \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\left(\frac{x}{\alpha_1} \right)^{\beta_1} \right)^{(1/\beta_1+1-1)} \frac{x^{(\beta_1-1)}}{(\alpha_1)^{\beta_1-1}} dx \tag{C.8}$$

or,

$$\Gamma(1/\beta_1 + 1) = \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\frac{x}{\alpha_1} \right)^{\beta_1 (1/\beta_1 + 1 - 1)} \frac{\beta_1}{(\alpha_1)^{\beta_1}} x^{(\beta_1 - 1)} dx = \frac{\beta_1}{\alpha_1} \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\frac{x}{\alpha_1} \right)^{\beta_1} dx \quad (C.9)$$

Alors,

$$P_2 = \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \frac{\Gamma(1/\beta_1)}{\beta_1} \frac{\alpha_1}{\beta_1} = \frac{1}{\beta_1} \quad (C.10)$$

Si on essaie d'assimiler P_3 à la formule suivante : $\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt$ avec $z > 0$, l'expression

P_3 s'écrit de cette façon :

$$\begin{aligned} P_3 &= \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\frac{x}{\alpha_2} \right)^{\beta_2} dx \\ &= \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\frac{x}{\alpha_2} \right)^{\beta_2} dx \\ &= \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \int_0^{\infty} e^{-(x/\alpha_1)^{\beta_1}} \left(\frac{x}{\alpha_1} \right)^{\beta_1} \left(\frac{\beta_2 + 1}{\beta_1} - 1 \right) \frac{(\alpha_1)^{(\beta_2 + 1)} \beta_1 x^{\beta_1 - 1}}{(\alpha_1)^{\beta_1} \beta_1 (\alpha_2)^{\beta_2}} dx \\ &= \frac{\beta_1}{\alpha_1 \Gamma(1/\beta_1)} \frac{(\alpha_1)^{(\beta_2 + 1)} \Gamma((\beta_2 + 1)/\beta_1)}{\beta_1 (\alpha_2)^{\beta_2}} \\ &= \left(\frac{\alpha_1}{\alpha_2} \right)^{\beta_2} \frac{\Gamma((\beta_2 + 1)/\beta_1)}{\Gamma(1/\beta_1)} \end{aligned} \quad (C.11)$$

Finalement, l'équation (C.5) s'écrit de la manière suivante :

$$\begin{aligned} D(f(., \alpha_1, \beta_1); f(., \alpha_2, \beta_2)) &= \log \left(\frac{\beta_1 \alpha_2 \tau(1/\beta_2)}{\beta_2 \alpha_1 \tau(1/\beta_1)} \right) + \\ &+ \left(\frac{\alpha_1}{\alpha_2} \right)^{\beta_2} \frac{\tau((\beta_2 + 1)/\beta_1)}{\tau(1/\beta_1)} - \frac{1}{\beta_1} \end{aligned} \quad (C.12)$$

Annexe D

Les tableaux D.1 et D.2 présentent l'application des formules de correction de décalage sur les paquets (au total 62 paquets) de chaque niveau d'échelle ($j = 1, \dots, 5$). L'algorithme de détection du maximum de vraisemblance est appliqué sur les 1000 signaux de la base SIM3. Les moyennes des instants de détection ont été calculées sur chaque paquet. Une première correction simple a été faite par multiplication par 2^j , puis la correction de décalage est calculée. La différence entre l'instant réel de détection 2000 et l'instant obtenu après correction a été calculée.

j=1	2005.9	4.5500	2001.4	1.35	j=5	2138.9	141.0500	1997.9	-2.15
	2005.9	4.4500	2001.5	1.45		2147.4	140.9500	2006.5	6.45
j=2	2014.8	13.6500	2001.2	1.15	2146.8	140.8500	2006	5.95	
	2015.8	13.5500	2002.3	2.25	2145.2	140.7500	2004.5	4.45	
	2015.6	13.4500	2002.2	2.15	2146.1	140.6500	2005.5	5.45	
	2015.5	13.3500	2002.2	2.15	2149.3	140.5500	2008.8	8.75	
j=3	2031.7	31.8500	1999.9	-0.15	2147.6	140.4500	2007.2	7.15	
	2034	31.7500	2002.3	2.25	2141.5	140.3500	2001.2	1.15	
	2034	31.6500	2002.4	2.35	2143.4	140.2500	2003.2	3.15	
	2033.2	31.5500	2001.7	1.65	2148.4	140.1500	2008.3	8.25	
	2032.8	31.4500	2001.4	1.35	2147.8	140.0500	2007.8	7.75	
	2034.1	31.3500	2002.8	2.75	2144.4	139.9500	2004.5	4.45	
	2033.3	31.2500	2002.1	2.05	2144.8	139.8500	2005	4.95	
	2033.1	31.1500	2002	1.95	2150.9	139.7500	2011.2	11.15	
j=4	2066.7	68.2500	1998.5	-1.55	2149.5	139.6500	2009.9	9.85	
	2068.5	68.1500	2000.4	0.35	2144.8	139.5500	2005.3	5.25	
	2073.5	68.0500	2005.5	5.45	2140.9	139.4500	2001.5	1.45	
	2070.7	67.9500	2002.8	2.75	2147.7	139.3500	2008.4	8.35	

2069.6	67.8500	2001.8	1.75	2153.4	139.2500	2014.2	14.15
2071.6	67.7500	2003.9	3.85	2144.9	139.1500	2005.8	5.75
2072.4	67.6500	2004.8	4.75	2144.8	139.0500	2005.8	5.75
2069.1	67.5500	2001.6	1.55	2152.3	138.9500	2013.4	13.35
2067.1	67.4500	1999.7	-0.35	2148.3	138.8500	2009.5	9.45
2071	67.3500	2003.7	3.65	2143.3	138.7500	2004.6	4.55
2073.8	67.2500	2006.6	6.55	2143.3	138.6500	2004.7	4.65
2070.8	67.1500	2003.7	3.65	2148.4	138.5500	2009.9	9.85
2069.6	67.0500	2002.6	2.55	2150.6	138.4500	2012.2	12.15
2070.9	66.9500	2004	3.95	2143.7	138.3500	2005.4	5.35
2073.7	66.8500	2006.9	6.85	2146.9	138.2500	2008.7	8.65
2068.6	66.7500	2001.9	1.85	2150.9	138.1500	2012.8	12.75
				2145.3	138.0500	2007.3	7.25
				2142.9	137.9500	2005	4.95

Tableau D.1. Pour les 62 paquets d'ondelettes : colonne 1 : numéro de l'échelle, colonne 2 : instant de changement détecté, colonne 3 : correction simple liée à la décimation, colonne 4 : correction de décalage, colonne 5 : écart à la vraie valeur de détection (2000). Résultats dans le cas de l'ondelette Symlet no.5.

j=1	2019.9	10.0333	2009.8667	9.8667	j=5	2592.8	311.0323	2281.7677	281.7677
	2011.3	18.9667	1992.3333	-7.6667		2447.8	319.9657	2127.8343	127.8343
j=2	2058.5	30.0999	2028.4001	28.4001		2524.7	328.8991	2195.8009	195.8009
	2040.1	39.0333	2001.0667	1.0667		2379.6	337.8325	2041.7675	41.7675
	2049.3	47.9667	2001.3333	1.3333		2559.5	346.7659	2212.7341	212.7341
	2031.6	56.9001	1974.6999	-25.3		2414	355.6993	2058.3007	58.3007
	2135.8	70.2331	2065.5669	65.5669		2490.1	364.6327	2125.4673	125.4673
j=3	2098.9	79.1665	2019.7335	19.7335		2344.4	373.5661	1970.8339	-29.1661
	2116.1	88.0999	2028.0001	28.0001		2577.5	382.4995	2195.0005	195.0005
	2080.4	97.0333	1983.3667	-16.633		2431.6	391.4329	2040.1671	40.1671
	2126.8	105.9667	2020.8333	20.8333		2504.4	400.3663	2104.0337	104.0337
	2090.1	114.9001	1975.1999	-24.8		2363.5	409.2997	1954.2003	-45.7997
	2108.9	123.8335	1985.0665	-14.933		2541.2	418.2331	2122.9669	122.9669
	2071.6	132.7669	1938.8331	-61.167		2397.7	427.1665	1970.5335	-29.4665
j=4	2288.2	150.4995	2137.7005	137.701		2469	436.0999	2032.9001	32.9001

2214.4	159.4329	2054.9671	54.9671	2327.3	445.0333	1882.2667	-117.7333
2252.2	168.3663	2083.8337	83.8337	2587.2	453.9667	2133.2333	133.2333
2182.2	177.2997	2004.9003	4.9003	2441.9	462.9001	1978.9999	-21.0001
2270.2	186.2331	2083.9669	83.9669	2514.9	471.8335	2043.0665	43.0665
2198.2	195.1665	2003.0335	3.0335	2370	480.7669	1889.2331	-110.7669
2235.1	204.0999	2031.0001	31.0001	2554.1	489.7003	2064.3997	64.3997
2161.3	213.0333	1948.2667	-51.733	2404.8	498.6337	1906.1663	-93.8337
2279.1	221.9667	2057.1333	57.1333	2481.6	507.5671	1974.0329	-25.9671
2206.9	230.9001	1975.9999	-24	2335.4	516.5005	1818.8995	-181.1005
2243.4	239.8335	2003.5665	3.5665	2568	525.4339	2042.5661	42.5661
2173.8	248.7669	1925.0331	-74.967	2424.6	534.3673	1890.2327	-109.7673
2262.9	257.7003	2005.1997	5.1997	2496.8	543.3007	1953.4993	-46.5007
2190.7	266.6337	1924.0663	-75.934	2352.9	552.2341	1800.6659	-199.3341
2225.2	275.5671	1949.6329	-50.367	2530	561.1675	1968.8325	-31.1675
2151.9	284.5005	1867.3995	-132.6	2385.8	570.1009	1815.6991	-184.3009
				2461.1	579.0343	1882.0657	-117.9343
				2314.9	587.9677	1726.9323	-273.0677

Tableau D.2. Résultats dans le cas de l'ondelette Coiflet no.5.

Annexe E

Le problème de minimisation d'une fonction objective avec contraintes linéaires se résout en se basant sur plusieurs principes, parmi lesquels ceux cités ci dessous.

Principe de Fermat (1638)

Les points qui minimisent ou maximisent une fonction dérivable annulent sa dérivée. Ils sont appelés *points stationnaires*.

Principe de Lagrange (1788)

Pour résoudre un problème d'optimisation sous contrainte, il suffit de rechercher un point stationnaire z_0 du lagrangien $L(z, \alpha)$ de la fonction g à optimiser et les fonctions C_i^g exprimant les contraintes :

$$L(z, \alpha) = g(z) + \sum_{i=1}^m \alpha_i C_i^g(z)$$

où les $\alpha_i = (\alpha_1, \dots, \alpha_m)$ sont des constantes appelés *coefficients de Lagrange*.

Principe de Kuhn-Tucker (1951)

Avec des fonctions g et C_i^g convexes, il est toujours possible de trouver un point selle (z_0, α^*) qui vérifie

$$\min_z L(z, \alpha^*) = L(z_0, \alpha^*) = \max_{\alpha \geq 0} L(z_0, \alpha)$$

Les références [Cia94], [Fle87] contiennent plus de détails sur ces trois principes.

Annexe F

La méthode SVM multiclasse est appliquée aux événements de l'EMG utérin. Chaque tableau est calculé pour un σ donné et pour $C = [0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000, \text{et Inf}]$. Les tableaux F.1 à F.8 utilisent uniquement les caractéristiques calculées à partir des paquets 15, 16 et 7.

$\sigma = 0.1$	CT	Alv	MAF	LDBF	bruit
C=0.0001	0.64	0.54	0.6200	0.8200	0.52
0.001	0.64	0.54	0.62	0.82	0.52
0.01	0.64	0.54	0.62	0.82	0.52
0.1	0.64	0.54	0.62	0.82	0.52
1	0.64	0.54	0.62	0.78	0.52
10	0.64	0.54	0.62	0.76	0.52
100	0.64	0.54	0.64	0.72	0.52
1000	0.64	0.54	0.64	0.72	0.52
10000	0.64	0.54	0.64	0.72	0.52
Inf	0.64	0.54	0.64	0.72	0.52

Tableau F.1. Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.1$.

$\sigma = 0.2$	CT	Alv	MAF	LDBF	bruit
C=0.0001	0.62	0.52	0.66	0.84	0.52
0.001	0.62	0.52	0.66	0.84	0.52
0.01	0.62	0.52	0.66	0.84	0.52
0.1	0.62	0.52	0.66	0.84	0.52
1	0.62	0.52	0.66	0.80	0.5
10	0.64	0.56	0.62	0.80	0.5
100	0.64	0.56	0.62	0.76	0.5
1000	0.64	0.56	0.62	0.72	0.5
10000	0.64	0.56	0.62	0.72	0.5
Inf	0.64	0.56	0.62	0.72	0.5

Tableau F.2. Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.2$.

$\sigma = 0.4$	CT	Alv	MAF	LDBF	bruit
C=0.0001	0.6	0.3	0.68	0.8400	0.4000
0.001	0.6	0.3	0.68	0.84	0.40
0.01	0.6	0.3	0.68	0.84	0.40
0.1	0.6	0.3	0.68	0.84	0.40
1	0.6	0.42	0.70	0.80	0.52
10	0.64	0.54	0.62	0.80	0.54
100	0.64	0.58	0.60	0.86	0.52
1000	0.62	0.58	0.60	0.82	0.52
10000	0.62	0.58	0.62	0.78	0.52
Inf	0.62	0.58	0.62	0.78	0.52

Tableau F.3 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.4$.

$\sigma = 0.6$	CT	Alv	MAF	LDBF	bruit
C=0.0001	0.54	0.2600	0.6600	0.82	0.22
0.001	0.54	0.26	0.66	0.82	0.22
0.01	0.54	0.26	0.66	0.82	0.22
0.1	0.56	0.26	0.66	0.82	0.22
1	0.6	0.34	0.68	0.84	0.56
10	0.66	0.5	0.70	0.80	0.68
100	0.66	0.54	0.64	0.82	0.54
1000	0.72	0.54	0.68	0.84	0.56
10000	0.76	0.54	0.70	0.80	0.56
Inf	0.76	0.54	0.66	0.80	0.56

Tableau F.4 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.6$.

$\sigma = 0.8$	CT	Alv	MAF	LDBF	bruit
C=0.0001	0.4	0.1800	0.5200	0.74	0.1600
0.001	0.4	0.18	0.52	0.74	0.16
0.01	0.4	0.18	0.52	0.74	0.16
0.1	0.46	0.20	0.58	0.82	0.16
1	0.6	0.24	0.68	0.82	0.54
10	0.62	0.46	0.74	0.82	0.72
100	0.70	0.56	0.70	0.82	0.58
1000	0.74	0.56	0.62	0.86	0.60
10000	0.72	0.58	0.64	0.86	0.60
Inf	0.68	0.56	0.64	0.78	0.60

Tableau F.5 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.8$.

$\sigma = 1$	CT	Alv	MAF	LDBF	bruit
C=0.0001	0.1	0.1800	0.4800	0.6600	0.1600
0.001	0.1	0.18	0.48	0.66	0.16
0.01	0.1	0.18	0.48	0.66	0.16
0.1	0.3	0.20	0.56	0.74	0.16
1	0.58	0.24	0.66	0.80	0.56
10	0.6	0.46	0.72	0.84	0.68
100	0.66	0.56	0.68	0.84	0.68
1000	0.76	0.56	0.64	0.84	0.64
10000	0.78	0.58	0.64	0.86	0.60
Inf	0.70	0.56	0.64	0.88	0.60

Tableau F.6 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 1$.

$\sigma = 5$	CT	Alv	MAF	LDBF	bruit
C=0.0001	0.	0	0	0	0
0.001	0	0	0	0	0
0.01	0	0	0	0	0
0.1	0	0	0	0	0
1	0	0	0.50	0.60	0.10
10	0	0.2	0.72	0.60	0.72
100	0.7	0.28	0.74	0.68	0.82
1000	0.7	0.28	0.78	0.82	0.78
10000	0.7	0.28	0.80	0.84	0.76
Inf	0.82	0.48	0.80	0.84	0.70

Tableau F.7 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 5$.

$\sigma = 10$	CT	Alv	MAF	LDBF	bruit
C=0.0001	0.1	0	0	0	0
0.001	0	0	0	0	0
0.01	0	0	0	0	0
0.1	0	0	0	0	0
1	0	0	0	0	0
10	0	0.04	0.56	0.6	0.46
100	0.02	0.16	0.70	0.60	0.76
1000	0.74	0.24	0.78	0.60	0.78
10000	0.66	0.28	0.76	0.76	0.78
Inf	0.72	0.64	0.80	0.88	0.70

Tableau F.8 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 10$.

Les tableaux E.9 à E.16 utilisent les caractéristiques calculées à partir des paquets 15, 16 et 7, en plus que la longueur des événements.

$\sigma = 0.1$	CT	Alv	MAF	LDBF	bruit
C=0.0001	0.2600	0.4800	0.84	0.26	0.86
0.001	0.26	0.48	0.84	0.26	0.86
0.01	0.26	0.48	0.84	0.26	0.86
0.1	0.26	0.48	0.84	0.26	0.86
1	0.26	0.50	0.84	0.26	0.86
10	0.26	0.50	0.84	0.26	0.86
100	0.26	0.50	0.84	0.26	0.86
1000	0.26	0.50	0.84	0.26	0.86
10000	0.26	0.50	0.84	0.26	0.86
Inf	0.26	0.50	0.84	0.26	0.86

Tableau F.9 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.1$.

$\sigma = 0.2$	CT	Alv	MAF	LDBF	Bruit
C=0.0001	0.4200	0.84	0.94	0.26	0.88
0.001	0.42	0.84	0.94	0.26	0.88
0.01	0.42	0.84	0.94	0.26	0.88
0.1	0.42	0.84	0.94	0.26	0.88
1	0.42	0.84	0.94	0.26	0.88
10	0.42	0.84	0.94	0.26	0.88
100	0.42	0.84	0.94	0.26	0.88
1000	0.42	0.84	0.94	0.26	0.88
10000	0.42	0.84	0.94	0.26	0.88
Inf	0.42	0.84	0.94	0.26	0.88

Tableau F.10 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.2$.

$\sigma = 0.4$	CT	Alv	MAF	LDBF	Bruit
C=0.0001	0.5400	0.84	0.94	0.38	0.88
0.001	0.54	0.84	0.94	0.38	0.88
0.01	0.54	0.84	0.94	0.38	0.88
0.1	0.54	0.84	0.94	0.38	0.88
1	0.54	0.84	0.94	0.38	0.88
10	0.54	0.84	0.94	0.38	0.88
100	0.54	0.84	0.94	0.38	0.88
1000	0.54	0.84	0.94	0.38	0.88
10000	0.54	0.84	0.94	0.38	0.88
Inf	0.54	0.84	0.94	0.38	0.88

Tableau F.11 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.4$.

$\sigma = 0.6$	CT	Alv	MAF	LDBF	Bruit
C=0.0001	0.58	0.84	0.94	0.48	0.9
0.001	0.62	0.84	0.94	0.48	0.9
0.01	0.62	0.84	0.94	0.48	0.9
0.1	0.62	0.84	0.94	0.48	0.9
1	0.62	0.84	0.94	0.48	0.9
10	0.62	0.84	0.94	0.48	0.9
100	0.62	0.84	0.94	0.48	0.9
1000	0.62	0.84	0.94	0.48	0.9
10000	0.62	0.84	0.94	0.48	0.9
Inf	0.62	0.84	0.94	0.48	0.9

Tableau F.12 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.6$.

$\sigma = 0.8$	CT	Alv	MAF	LDBF	Bruit
C=0.0001	0.66	0.8400	0.9400	0.54	0.9
0.001	0.66	0.84	0.94	0.54	0.9
0.01	0.66	0.84	0.94	0.54	0.9
0.1	0.66	0.84	0.94	0.54	0.9
1	0.66	0.84	0.94	0.54	0.9
10	0.66	0.84	0.94	0.54	0.9
100	0.66	0.84	0.94	0.54	0.9
1000	0.66	0.84	0.94	0.54	0.9
10000	0.66	0.84	0.94	0.54	0.9
Inf	0.66	0.84	0.94	0.54	0.9

Tableau F.13 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 0.8$.

$\sigma = 1$	CT	Alv	MAF	LDBF	Bruit
C=0.0001	0.70	0.84	0.94	0.58	0.9
0.001	0.70	0.84	0.94	0.58	0.9
0.01	0.70	0.84	0.94	0.58	0.9
0.1	0.70	0.84	0.94	0.58	0.9
1	0.70	0.84	0.94	0.58	0.9
10	0.70	0.84	0.94	0.58	0.9
100	0.70	0.84	0.94	0.58	0.9
1000	0.70	0.84	0.94	0.58	0.9
10000	0.70	0.84	0.94	0.58	0.9
Inf	0.70	0.84	0.94	0.58	0.9

Tableau F.14 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 1$.

$\sigma = 5$	CT	Alv	MAF	LDBF	Bruit
C=0.0001	0.70	0.86	0.96	0.94	0.9
0.001	0.70	0.86	0.96	0.94	0.90
0.01	0.70	0.86	0.96	0.94	0.90
0.1	0.70	0.86	0.96	0.94	0.90
1	0.82	0.86	0.96	0.94	0.90
10	0.82	0.86	0.96	0.940	0.90
100	0.84	0.86	0.96	0.94	0.90
1000	0.86	0.86	0.96	0.94	0.90
10000	0.86	0.86	0.96	0.94	0.90
Inf	0.88	0.86	0.96	0.94	0.9

Tableau F.15 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 5$.

$\sigma = 10$	CT	Alv	MAF	LDBF	Bruit
C=0.0001	0.70	0.88	0.96	0.98	0.78
0.001	0.70	0.82	0.96	0.98	0.78
0.01	0.70	0.82	0.96	0.98	0.78
0.1	0.70	0.82	0.96	0.98	0.78
1	0.70	0.820	0.960	0.98	0.78
10	0.70	0.82	0.96	0.98	0.78
100	0.80	0.820	0.96	0.98	0.78
1000	0.80	0.82	0.96	0.98	0.90
10000	0.80	0.82	0.96	0.98	0.90
100000	0.80	0.82	0.96	0.98	0.90
Inf	0.80	0.82	0.96	0.98	0.90

Tableau F.16 Taux de bonne classification de l'ensemble test de chacun des événements de l'EMG utérin pour $\sigma = 10$.

Bibliographie

- [Abr70] Abramowitz M., Stegun I. A., *Handbook of Mathematical Tables*. New York, Dover, 1970.
- [Abr94] Abry P., "Multiresolution transient detection", *Proc. of IEEE TFTS conference*, Philadeiphia, USA, October 1994.
- [Abr97] Abry P., *Ondelettes et turbulences*, Diderot Editeur, Arts et sciences, Paris, 1997.
- [Akh05] Akhbardeh A., Koivuluoma M., "Ballistocardiogram diagnosis using neural networks and shift-invariant daubechies wavelet transform", *13th European Signal Conference EUSIPCO*, Antalya Turkey, 4-8 September 2005.
- [Alb03] Albanesi M. G., Guerrini F., "An HVS-based adaptive coder for perceptually lossy image compression", *Pattern Recognition*, vol. 36, no. 4, pp. 997-1007, 2003.
- [Ant01] Antoniadis A., Bigot J., Sapatinas T., "Wavelet estimators in nonparametric regression : a comparative simulation study", *Journal of Statistical Software*, vol. 6, no. 6, pp. 1-83, 2001. <http://www-lmc.imag.fr/lmc-sms/Anestis.Antoniadis/HTTP/publis-anto.html>
- [Bas89] Basseville M., "Distances measures for signal processing and pattern recognition", *Signal processing*, vol. 19, pp. 349-369, 1989.
- [Bas93] Basseville M., Nikiforov I., *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [Boa90] Boashash B., Oshea P., "A methodology for detection and classification of some underwater acoustic signals using time frequency analysis techniques", *IEEE Trans. ASSP*, vol. 38, no. 11, pp. 1829-1841, 1990.
- [Bod31] Bode O., *Das elektrohysterogram*, Arch. Gynaekol., vol. 153, pp. 123, 1931.
- [Bor87] Borovkov A., *Statistique mathématique*, Mir, Moscou, 1987.

- [Bra83] Brandt A.V., "Detecting and estimating parameters jumps using ladder algorithms and likelihood ratio test", *Proc. of ICASSP*, Boston, MA, pp. 1017-1020, 1983.
- [Bri92] D. Brie, "Méthodes statistiques de détection de rupture de modèle. Application au traitement du signal issu d'un capteur à courants de Foucault", *Doctorat de l'université, spécialité automatique*, Université de Nancy I, France, 1992.
- [Cal93] Caleydro-Brica R., Poseiro J.J., "Physiology of uterine contraction", *Clin Obstet Gynecol*, vol. 3, pp. 1636-1653, 1993.
- [Car93] Carrault G., Bellanger J.J., Badier J.M., "Segmentation vectorielle des signaux EEG", *Proc. XIV ème colloque GRETSI*, pp. 767-770, Juan-Les-Pins, France, 1993.
- [Che04a] Chendeb M., Khalil M., Duchêne J., "Wavelet Based Method for Detection: Application in Proprioceptive Rehabilitation", *26th Annual International Conference of the IEEE EMBS Engineering in Medicine and Biology Society*, San Francisco, USA, 1-5 September 2004.
- [Che04b] Chendeb M., Khalil M., Duchêne J., "Wavelet-based method for detection of events in long duration signals", *5th EUROSIM Congress on Modeling and Simulation*, ESIEE Paris, Marne la Vallée, France, 6-10 September 2004.
- [Che05a] Chendeb M., Khalil M., Duchêne J., "New Approach for Detection using Wavelet Coefficients", *ICITA'2005 3rd International conference on Information Technology and Applications*, Sydney, Australie, 4-7 July 2005.
- [Che05b] Chendeb M., Khalil M., Duchêne J., "The use of wavelet packets for event detection", *13th European Signal Conference EUSIPCO*, Antalya Turkey, 4-8 September 2005.
- [Che05c] Chendeb M., Khalil M., Duchêne J., "The use of wavelet packets for detection and identification of events in uterine EMG", *3rd European Medical & biological engineering conference, IFMBE European conference on biomedical engineering*, Prague, Czech republic, 20 -25 November 2005.
- [Che06] Chendeb M., Khalil M., Duchêne J., "Methodology of Wavelet Packet for Event Detection", accepted in *Signal Processing*, January 2006.
- [Cia94] Ciarlet P., *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, nouvelle édition, 1994.
- [Coi92] Coifman R.R, Wickerhauser M.V., "Entropy_based algorithms for best basis selection", *IEEE Trans. Information Theory*, vol. 38, no. 2, pp. 713-718, March 1992.
- [Cov91] Cover T. M., Thomas J. A., *Elements of Information Theory*, New York, Wiley, 1991.

- [Dau01] Daudet L., Molla S., Torrèsani B., "Transient detection and encoding using wavelet coefficient trees", *Proceedings of the GRETSI Conference*, Toulouse, France, September 2001.
- [Dau90] Daubechies I, "The wavelet transform time-frequency localization and signal analysis", *IEEE Trans. Information Theory*, vol. 36, no.5, pp. 961-1005, 1990.
- [Dev95a] Devedeux D., "Evaluation quantitative de certaines caractéristiques de distributions temps fréquence: application à l'EMG utérin", Thèse soutenue à l'Université de Technologie de Compiègne, 1995.
- [Dev95b] Devedeux D., Duchêne J., Germain G. et al., "Evolution Temporelle des principales composantes spectrales de l'électromyogramme utérin : étude expérimentale", *Innov. Techn. Biol. Med.*, vol. 16, no. 4, pp. 496-480, 1995.
- [Din01] Dinh H. A. N., Kumar D. K., et al., "Wavelets for QRS Detection", *23th Annual International Conference of the IEEE EMBS Engineering in Medicine and Biology Society*, vol. 2, pp. 1883-1887, October 2001.
- [Do02] Do M. N., Vetterli M., "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance" *Proc. IEEE Trans Image Processing*, vol. 11, no. 2, pp. 146-158, 2002.
- [Don94] D. Donoho and I. Johnstone, "Ideal denoising in an orthogonal basis chosen from a library of bases", *Comptes Rendus Aca. Sci Paris*, no. 319, pp. 1317-1322, October 1994.
- [Don95a] Donoho D., "Denoising via soft-thresholding", *IEEE Trans Inform. Theory*, vol. 41, no.3, pp. 613-627, 1995.
- [Don95b] Donoho D., Johnstone M.L., "Adapting to unknown smoothness via wavelet shrinkage", *Journal of American statistical association*, vol.90, no. 432, pp. 1200-1224, 1995.
- [Dua05] Duan K. B., Keerthi S. S., "Which Is the Best Multiclass SVM Method? An Empirical Study", *Multiple Classifiers Systems*, pp. 278-285, 2005.
- [Dub90] Dubuisson B., "*Diagnostic et reconnaissance des formes*", Editions Hermes 1990.
- [Duc93] Duchêne J., Goubel F., "Surface Electromyogram during voluntary contraction: processing tools and relation to physiological events", *Crit. Rev. in Biomed. Eng.*, vol. 21, no. 4, pp. 313-397, 1993.
- [Duc95] Duchêne J., Devedeux D., Mansour S., et al., "Analysing uterine EMG : Tracking instantaneous burst frequency", *IEEE Eng. In Medicine and Biology Magazine*, vol. 14, pp. 125-131, 1995.
- [Fal01] Falou E. W., Duchêne J, Grabish M., Hewson D., Langeron Y., Lino F., "Analysis of muscular fatigue in passengers sitting in a car seat", *Proc. of the*

International Society of Biomechanics, ISB XVIIIth, Zurich, Swiss, June 2001.

- [Fal02] Falou W., "Une approche de la segmentation dans des signaux de longue durée fortement bruités. Application en ergonomie", Thèse de Doctorat de l'Université de Technologie de Troyes, France, Décembre 2002.
- [Fle87] Fletcher R., "*Practical methods of optimization*", John Wiley, 2nd edition, 1987.
- [Fuk90] Fukunaga K., "*Introduction to statistical pattern recognition*", Academic Press, Inc. 1990.
- [Gov03] Govaert G, "Analyse des données", Hermes Lavoisier, Paris, 2003.
- [Gra95] Gramatikov B., Yi-chun S., Brinker J. A., Thakor N. V., "Multiresolution wavelet analysis of the body surface ECG before and after angioplasty", *Ann. Biomed. Eng.*, vol. 23. pp. 553-561, 1995.
- [Gun98] Gunn S.R., "Support Vector Machines for Classification and Regression", *Technical Report, School of Electronics and Computer Science*, University of Southampton, (Southampton, U.K.), 1998.
- [Had00] Hadjileontiadis L., Liatsos L., Marvrogianis C., Rokkas T., Panas S., "Enhancement of bowel sounds by wavelet-based filtering", *IEEE Trans. Biomedical Engineering*, vol. 47, no. 7, pp. 876-886, 2000.
- [Hig93] Higby K., Xenakis EMJ, Pauerstein CJ, "Do tocolytic agents stop preterm labor ? A critical and comprehensive review of efficacy and safety", *Am. J. Obstet. Gynecol.*, vol.168, no. 4, pp.1247-1256, 1993.
- [Hit98] Hitti H., Lucas M.F., "Wavelet-packet basis selection for abrupt changes detection in multicomponent signals" 9th European Signal Processing Conference EUSIPCO, Rhodes, Grèce, September 1998.
- [Hit99] Hitti. E., "Sélection d'un banc optimal de filtres à partir d'une décomposition en paquets d'ondelettes. Application à la détection de sauts de fréquences dans des signaux multicomposantes", Thèse de Doctorat, Automatique et Informatique, Ecole Centrale de Nantes, novembre 1999.
- [Hon58] Hon EHG, Davis CD, "Cutaneous and utérine electrical potentials in labor", *Obstet. Gynecol.*, vol. 12, no. 1, pp. 47-53, 1958.
- [Hop94] Hopper T, "Compression of grey-scale fingerprint images, In wavelet applications", *Proc. of SPIE, Orlando*, vol. 2242, p.180.
- [Hor96] Horovitz J, Guyon F, Roux D., Dubecq JP, "Accouchement prématuré", *Encycl. Med. Chir., Elsevier*, Paris, p. 14, 1996.
- [Inb84] Inbar G. F., Noujaim A. E., "On surface EMG spectral characterization and its application to diagnostic classification", *IEEE Trans. Biomedical Eng.*, vol. 31, no.9, pp. 597-604, 1984.

- [Joh93] Johnson P., "Suppression of preterm labour", current concepts, *Drugs*, 1993, vol. 45, pp 684-692, 1993.
- [Kad92a] Kadambe S., Boudreaux-Bartels G., "Application of the wavelet transform for pitch detection of speech signals". *IEEE Trans. Information Technology*, vol. 38, no. 2, pp 917-924, 1992.
- [Kad92b] Kadambe S., Murray R., Boudreaux-Barlets G.F., "The dyadic wavelet transform based QRS detector", *Proc. Conference Record of The Twenty-Sixth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 130-134, 1992.
- [Kad99] Kadambe S., Murray R., Boudreaux-Bartels G.F., "Wavelet transform-based QRS complex detector", *IEEE Trans. on Biomedical Engineering*, vol. 46, no. 7, pp. 838 – 848, July 1999.
- [Kal95] Kalayci T., Özdamar O., "Wavelet preprocessing for automated neural network detection of EEG spikes", *IEEE Engineering in Medicine and Biology Magazine*, vol. 13, pp. 160-166, 1995.
- [Kar95] Karrakchou M., Kunt M., "Multiscale analysis for singularity detection in pulmonary microvascular pressure transients", *Annals of biomedical engineering*, vol. 23, pp. 562-573, 1995.
- [Kay88] Kay S., "*Modern Spectral Estimation Theory & Application*", Englewood Cliffs, Prentice-Hall, NJ, USA, 1988.
- [Kay93] Kay S. M., "*Fundamentals of Statistical Signal Processing: Estimation Theory*", Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [Kha00] Khalil M., Duchêne J., "Uterine EMG Analysis: A dynamic approach for change detection and classification", *IEEE Trans. Biomed. Eng.*, vol. 47, no. 6, pp. 748-756, 2000.
- [Kha97] Khalil M., Duchêne J., Marque C., "Detection and isolation of multiple events in non-stationary signals by multiscale decomposition. Application to uterine EMG", *Intelligent Engineering systems through artificial neural networks ASME press*, USA, 1997.
- [kha99] Khalil M., Duchêne J., "Detection and classification of multiple events in piecewise stationary signals: Comparison between autogressive and multiscale approaches", *Signal Processing*, vol. 75, no. 3, pp. 239-251, 1999.
- [Kha99a] Khalil M., "Une approche de la détection et de la classification dans les signaux non stationnaires. Application à l'EMG utérin", *Thèse de Doctorat de l'Université de Technologie de Troyes*, France, Septembre 1999.
- [Kha02] Kharroubi J., "Etude de techniques de Classement "Machines à Vecteurs Supports" pour la vérification automatique du Locuteur", thèse de doctorat de l'Ecole Nationale Supérieure de Télécommunications, Juillet 2002.

- [Kri95] Krim H., Pesquet J.C., "Multiresolution analysis of a class of non stationary processes", *IEEE Trans. on information theory*, vol. 41, no. 4, pp.1011-1020, 1995.
- [Kul51] Kullback S., Leibler R. A., "On information and sufficiency", *Ann. Math. Statis.*, vol. 22, pp. 79-86, 1951.
- [Lea92] Learned R.E., Karl W.C., Willsky A. S., "Wavelet packet based transient signal classification", *Massachusetts Institute of Technology, Laboratory for Information and Decision Systems*, 1992.
- [Lea95] Learned R.E., Willsky A.S., "A wavelet packet approach to transient signal classification", *Applied and Computational Harmonic Analysis*, vol. 2, pp. 265-278, 1995.
- [Lee05] Lee K., Estivill-Castro V., "Feature extraction and gating techniques for ultrasonic shaft signal classification", *Applied Soft Computing, In Press, Corrected Proof*, August 2005.
- [Lem00] Leman H., Brosse V., Marque C., "Comparison between Wavelet Packets and Adaptive Filtering for the EHG Signal Denoising", *SCI ISAS*, 2000.
- [Lem95] Lemoine O., "Détection de signaux non stationnaires par représentation temps-fréquence", Thèse de doctorat à l'université de Nice-Sophia Antipolis, 1995.
- [Lem97] Leman H., Marque C., "Optimum denoising method for the uterine EHG", *Proc. of XIXth Annual International Conference of the IEEE EMBS Engineering in Medicine and Biology Society*, Chicago, USA, 28 Oct.-02 Nov. 1997.
- [Lem98] Leman H., Marque C., "Ridge extraction from the scalogram of the uterine electromyogram", *Proc. of IEEE TFTS*, 1998.
- [Let00] Letelier J.C., Weber P.P., "Spike sorting based on discrete wavelet transform coefficients", *Journal of Neuroscience Methods*, vol. 101, no. 2, pp. 93-106, September 2000.
- [Li05] Li D., Pedrycz W., Pizzi N.J., "Fuzzy wavelet packet based feature extraction method and its application to biomedical signal classification", *IEEE Trans. on Biomedical Engineering*, vol. 52, no. 6, pp. 1132- 1139, June 2005.
- [Li95] Li C. Zheng C., Tai C., "Detection of ECG characteristic points using wavelet transforms", *IEEE Trans. BME*, vol. 42, no. 1, pp 21-29, 1995.
- [Liu95] Liu W.Y., Magnin I.E., Gimenez G., "Un nouvel opérateur pour la détection de ruptures dans les signaux bruités", *Traitement du signal*, vol.12, no. 3, pp. 225-235, 1995.
- [Lop84] Lopes P., Germain G., Breart G., Reinato S., Le Houezec R., Sureau C., "Electromyographical study of uterine activity in the human during labor

- induced by prostaglandin F2 alpha", *Gynecol. Obstet. Invest.*, vol. 17, no.2, pp. 96-105, 1984.
- [Mal89a] Mallat S., "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pattern Recognit. Machine Intell PAMI*, vol. 11, no. 7, pp.674-693, July 1989.
- [Mal89b] Mallat S., "Multiresolution approximations and wavelet orthonormal bases of $L^2(R^n)$ ", *Trans. Am. Soc.*, vol. 315, no. 1, pp. 69-87, September 1989.
- [Mal92] Mallat S., Zhong S., "Characterisation of signals from multiscale edges", *IEEE Trans. IEEE Trans. Pattern Recognit. Machine Intell PAMI*, vol. 14, no. 7, pp. 711-733, 1992.
- [Mal99] Mallat S., "*A Wavelet Tour of Signal Processing*", Academic Press, San Diego, CA, 1999.
- [Man91] Mansour S., Duchêne J., Germain G., Marque C., "Utérine EHG: Experimental and mathematical détermination of the relationship between internal and external recordings", *Proc. 13th Annu. IEEE EMBS Conf.*, Orlando, USA, 1991
- [Mar86] Marque C., Duchene J., Leclercq S., Panczer J. et Chaumont J., "Uterine EHG processing for obstetrical monitoring", *IEEE transactions on biomedical engineering*, vol. 33, no. 12, pp. 1182-1187, 1986.
- [Mar87] Marple S. L., "*Digital spectral analysis with applications*", Prentice-Hall, Englewood Cliffs, NJ, USA, 1987.
- [Mar95] Marque C., Gondry J., Rossi J, Baaklini N. et Duchêne J., "Surveillance des grossesses à risque par électromyographie utérine", *revue européenne de technologie biomédicale*, vol. 17, no. 1, pp. 25-31, 1995.
- [May98] Mayoraz E., Alpaydin, E., "Support Vector Machines for Multi-Class Classification", *Technical Report IDIAP*, 1998.
- [McI92] McLachlan, G.J., "*Discriminant analysis and Statistical Pattern Recognition*", Wiley, New York, 1992.
- [Mer03] Merlo A., Farina D., Merletti R., "Fast and reliable technique for muscle activity detection from surface EMG signals", *IEEE Trans. on Biomedical Engineering*, vol. 50, no. 3, March 2003.
- [Mer99a] Merletti R., "The traditional Fourier approach to EMG analysis during voluntary or evoked sustained contractions", *SENIAM- Deliverable 7, State of the art on signal processing methods for SEMG*, pp. 65-78, 1999.
- [Mey90] Meyer Y., "*Ondelettes et Opérateurs*", tome 1, Hermann, Paris, 1990.
- [Mis04] Misiti M., Misiti Y., Oppenheim G., Poggi J.M., "*Wavelet Toolbox, For Use with MATLAB*", User's Guide Version 3, The MathWorks, Inc. 3 Apple Hill Drive, Natick, MA 01760-2098, 2004.

- [Mol02] Molla S., Torrèsani B., "Arbres de Markov Cachés de Coefficients d'Ondelettes pour la Détection de Transitoires dans les Signaux Audiophoniques", *Colloque Autosimilarité et Applications, Clermont-Ferrand, 27-30 May 2002*.
- [Mou99] Moulin P., Liu J., "Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors", *IEEE Trans. Inform. Theory*, vol. 45, pp. 909–919, 1999.
- [Nag88] Nageotte M., Dorchester W., Porto M., Keegan KA., Freeman R.K., "Quantitation of uterine activity preceding preterm, term, and postterm labor", *Am. J. Obstet. Gynecol.*, vol. 158, p. 1254-1259, 1988.
- [New87] Newman R.B., Gill P.J, Campion S., Katz M. "Anterpartum ambulatory to codynamometry: the significance of low- amplitude, high frequency contractions", *Obstet Gynecol*, vol. 70, pp 701-705, 1987.
- [Obr88a] Obrecht R.A., Su H. Y. "Three acoustic label lings for phoneme based continuous speech recognition", *Proc. Speech'88, Edinburgh*, pp. 943-950, 1988.
- [Obr88b] Obrecht R.A., "New statistical approach for the automatic segmentation of continuous speech signals", *IEEE Trans. Acoustic Speech Signal Process*, vol. 36, no.1, pp. 29-40, 1988.
- [Orf90] Orfandis S.J., "*Optimum Signal Processing an Introduction*", McGraw-Hill international editions, Singapore, 1990.
- [Pit99] Pittner S., Kamarthi S.V., "Feature extraction from wavelet coefficients for pattern recognition tasks", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 21, no. 1, pp. 83.88, 1999.
- [Pri94] Priebe R.D., Baugh K.W., "Wavelet based detectors", *Proc of IEEE Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, vol.4, pp. 105-108, Adelaide, Australia, April 19-22 1994.
- [Raj02] Rajpoot N.M., "Texture Classification Using Discriminant Wavelet Packet Subbands", *Proc. 45th IEEE Midwest Symposium on Circuits and Systems (MWSCAS'02)*, Tulsa, Oklahoma (USA), August 2002.
- [Raj04] Rajpoot K.M, Rajpoot M.N., "Wavelets and Support Vector Machines for Texture Classification," *Proc. of 8th IEEE-INMIC*, 2004.
- [Rav98] Ravier P., "Détection de transitoires par ondelettes adaptées. Critères d'adaptation fondés sur les statistiques d'ordre supérieur", *Doctorat de l'INPG Institut National Polytechnique de Grenoble, spécialité signal, image, parole*, France, 1998.
- [Ran01] Ranta R., Heinrich C., Louis-Dorr V., Wolf D., Guillemin F., "Wavelet-based bowel sounds denoising, segmentation and characterization", *Proceedings of IEEE EMBS Conference, Istanbul, Turkey, October 2001*.

- [Ran03] Ranta R., "Traitement et analyse de signaux sonores physiologiques. Application à la phonoentérogaphie", *Doctorat de l'Institut National Polytechnique de Lorraine, Spécialité Automatique et Traitement du Signal*, Décembre 2003.
- [Rav01] Ravier P., Amblard P.O., "Wavelet packets and de-noising based on higher-order-statistics for transient detection", *Signal Processing*, vol. 81, no. 9, pp. 1909–1926, 20 April 2001.
- [Rio93] Rioul O., "Ondelettes régulières: application à la compression d'images fixes", Thèse soutenue à l'ENST, 1993.
- [Ros62] Rosenblatt F, "*Principles of neurodynamics*", Spartan Press, East Lansing, MI, 1962.
- [Ruf88] Ruf H., Conte M., Franquelbalme JP, "L'accouchement prématuré", *Encycl. Med. Chir., Elsevier*, pp.12, Paris, 1988.
- [Sai02] Saito N., Coifman R., Geshwind F. B., Warner F., "Discriminant Feature Extraction Using Empirical Probability Density Estimation and a Local Basis Library", *Pattern Recognition*, vol. 35, pp.2841-2852, 2002.
- [Sai94] Saito N., Coifman R. R., "Local Discriminant Bases", *Proc. SPIE Mathematical Imaging: Wavelet Applications in Signal and Image Processing*, vol. 2303, pp. 65-75, July 1994.
- [Sai95] Saito N., Coifman R. R., "Local discriminant bases and their applications", *Journal Mathematical Imaging and Vision*, vol. 5, no. 4, pp. 337-358, 1995.
- [Sai97] Saito N., "Classification of geophysical acoustic waveforms using time-frequency atoms", *Proc. Computing Section of American. Statistical Association*, pp.322-327, 1997.
- [Sam69] Sammon J., "A Non Linear Mapping for Data Structure Analysis", *IEEE Transactions on Computing*, , no 18, pp. 401–409, 1969.
- [Sam95] Samar V.J., Swartz K. P., Raghuveer M.R., "Multiresolution analysis of event-related potentials by wavelet decomposition", *Brain and Cognition*, vol. 27, pp 398-438, 1995.
- [Sap90] Saporta G., "*analyse des données et statistiques*", éditions Technip, Paris, 1990.
- [Sch95] Schölkopf B., Burges C., Vapnik V., "Extracting support data for a given task", *In ICKDDM*, pp. 252-257, 1995.
- [Sen90] Senhadji L., Bellanger, G. Carrault, and J.L. Coatrieux, "Wavelet analysis of ECG signal", *Proc. Int. Cong. IEEE EMBS*, pp.811-812, Philadelphia, PA, 1990.

- [Sen93] Senhadji L., "Approaches multiresolution pour l'analyse des signaux non stationnaires", *Doctorat à l'université de Rennes, France*, 1993.
- [Sen95a] Senhadji L., Bellanger J.J., Carraut G., "Détection temps échelle d'événements paroxystiques intercritiques en électroencéphalogramme", *Traitement du Signal*, vol. 12, no. 4, pp 357-371, 1995.
- [Sen95bs] Senhadji L., Carraut G., Bellanger J.J., Passariella G. F., "Comparing wavelet transforms for recognizing cardiac patterns", *IEEE EMB Magazine*, vol. 13, pp.167-173, March 1995.
- [Sha95] Sharifi K., Leon-Garcia A., "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 52–56, 1995.
- [She01] Shen M., Sun L., Chan F.H. Y., "The Classification of Transient Time-Varying EEG Signals via Wavelet Packets Decomposition", *ICONIP 8th International Conference on Neural Information Processing*, Shanghai, China, 14-18. November 2001.
- [She92] Shensa M.J., "The Discrete Wavelet Transform: Wedding the à Trous and Mallat algorithms", *IEEE Trans. on information theory*, vol. 40, no. 10, pp. 2464-2482, October 1992.
- [She97] Sheng S., Yanda L., "Best Bases for Classification", *Technical Report of pattern Recognition and Signal Processing Lab.*, department of Automation, Tsinghua University, 1997.
- [Sir93] Siregar P., Coatrieux J. L., Mabo P., "How can deep knowledge be used in CCU monitoring", *IEEE Eng. Med. Biol.*, vol. 12, pp. 3397-3415, 1993.
- [Ste50] Steer C.M., Hertsh GJ., "Electrical activity of the human uterus in labor; the electrohysterograph", *Am. J. Obstet. Gynecol.*, vol. 59, no.1, pp. 25-40, 1950.
- [Sun05] Sun G., Dong X., Xu G., "Tumor tissue identification based on gene expression data using DWT feature extraction and PNN classifier", *Neurocomputing*, In Press, Corrected Proof, Available online 29 August 2005.
- [Sur56] Sureau C., "Etude de l'activité électrique de l'utérus au cours du travail", *Gynecol. Obstét.*, vol. 55, no. 2, pp. 153- 175, 1956.
- [Ter01] Terrien J., Catherine marque, "Mise en œuvre et optimisation de la segmentation et de la classification des événements contenus dans l'électrohysterogramme", *DEA sciences et technologie, laboratoire de biomécanique et génie biomédical, Université de Technologie de Compiègne*, 2001.
- [Tha93] Thakor N.V., Xin-rong G., Hanley D.F. "Multiresolution wavelet analysis of evoked potentials", *IEEE Trans. BME*, vol. 40, no. 11, pp 1085-1094, 1993.

- [Tho02] Thornburg H., "Detection and modeling of transients", *Center for Computer Research in Music and Acoustics (CCRMA)*, Department of Music, Stanford University, California, 2002. <http://www-crma.stanford.edu/~jos/tvar/tvar.html>.
- [Tir97] Tira Z., Duvaut P., Jousselin A. "Une Méthode de détection multiéchelle fondée sur le maximum de vraisemblance et les ondelettes multiples : définition et performances", *16^{ième} colloque GRETSI*, pp 1463-1466, 1997.
- [Tru98] Truchetet F., "*Ondelettes pour le signal numérique*", Editions HERMES, Paris, 1998.
- [Tur03] Turkoglu I., Ahmet Arslan A., Ilkay E., "An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks", *Computers in Biology and Medicine*, vol. 33, no.4, pp. 319-331, July 2003.
- [Tut87] Tuteur F.B., "Wavelet transformation in signal detection. Wavelets, Time-frequency methods and phase space", *Proc. of the international conference, Marseille, JM Combes, A Grossman, Ph. Tchamitchian eds. Springer-Verlag*, pp. 132-138, 14-18 December 1987.
- [Ube04] Übeyli E.D., Güler I., "Feature extraction from Doppler ultrasound signals for automated diagnostic systems", *Computers in Biology and Medicine*, 2004.
- [Uns96] Unser M., Aldroubi A., "A review of wavelets in biomedical applications", *Proc. IEEE*, vol. 84, no. 4, pp. 626-638, 1996.
- [Vap95] Vapnik V.N., "*The nature of statistical learning theory*", spring-Verlag, New York 1995.
- [Vap98] Vapnik V.N., "*Statistical learning theory*", John Wiley & Sons, Inc., N.Y., 1998.
- [Var89] Varanasi M.K., Aazhang B., "Parametric generalized Gaussian density estimation", *J. Acoust. Soc. Amer.* vol. 86, no. 4, pp. 1404–1415, 1989.
- [Voi00] Voisine M.L., Marque C., Leman H., "Réalisation d'un prototype d'appareillage pour le suivi à domicile de l'efficacité des contractions utérines pendant la grossesse", *Congrès mondial de télémédecine*, p.22, Toulouse, mars 2000.
- [Wal97] Walczak B., Massart D.L., "Noise suppression and signal compression using the wavelet packet transform", *Chemometrics Intell. Lab. Syst.*, vol 36, pp. 81-94, 1997.
- [Wei04] Wei Z., Yam L.H., Cheng L., "Detection of internal delamination in multi-layer composites using wavelet packets combined with modal parameter analysis", *Composite Structures*, vol. 64, pp. 377–387, 2004.
- [Wen96] Wendling F., "Mise en correspondance d'observations EEG de profondeurs pour la reconnaissance de signatures spatio-temporelles dans les crises

d'épilepsie", *Doctorat de l'université, mention traitement du signal et télécommunications*, Université de Rennes I, France, 1996.

- [Wic91] Wickerhauser M.V., "INRIA lectures on wavelet packet algorithms", *In Proc. Ondelettes et paquets d'ondes*, pp. 31-99, Rocquencourt, France, 17-21 June 1991.
- [Wol70] Wolfs G., Rottinghuis H., "Electrical and mechanical activity of the human uterus during labour", *Arch. Gynaekol.*, vol. 208, p. 373, 1970.
- [Wou99] Wouwer G.V., Scheunders P., Dyck D. V., "Statistical texture characterization from discrete wavelet representations", *IEEE Trans. Image Processing*, vol. 8, pp. 592–598, Apr. 1999.
- [Xu02] Xu P., Chan A. K., "Fast and robust neural network based wheel bearing fault detection with optimal wavelet features", *International Joint Conference on Neural Networks (IJCNN)*, 2002.
- [Yan05] Yan R. and Gao. R.X., "An efficient approach to machine health diagnosis based on harmonic wavelet packet transform", *Robotics and Computer-Integrated Manufacturing*, vol. 21, no. 4-5, pp. 291-301, 2005.
- [Yon05] Yong A., Hurley N., "Single Trial EEG Classification For Brain Computer using Wavelet Decomposition", *13th European Signal Conference EUSIPCO*, Antalya Turkey, 4-8 September 2005.