



HAL
open science

Classification des signaux EMG utérins afin de détecter les accouchements prématurés

M.O. Diab

► **To cite this version:**

M.O. Diab. Classification des signaux EMG utérins afin de détecter les accouchements prématurés. Sciences de l'ingénieur [physics]. Université de Technologie de Compiègne, 2007. Français. NNT : . tel-00410409

HAL Id: tel-00410409

<https://theses.hal.science/tel-00410409>

Submitted on 20 Aug 2009

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.

MEMOIRE PRESENTE POUR L'OBTENTION DU GRADE
DE DOCTEUR

Spécialité : Génie Biomédical

Présenté le 19 Décembre 2007 à
L'Université de Technologie de Compiègne (UTC)

par

Mohamad DIAB

**Classification des signaux EMG Utérins
Afin de Détecter Les Accouchements
Prématurés**

Guy Carrault	Professeur	LTSI	RENNES1	Rapporteur
Thierry Denoeux	Professeur	HEUDIASYC	UTC	Président du Jury
Mohamad Khalil	Maitre de conférences (HDR)	S2I	IUL	Co-directeur de Thèse
Dimitri Lefebvre	Professeur	GREAH	LE HAVRE	Rapporteur
Catherine Marque	Professeur	UMR-6600	UTC	Directeur de Thèse
Amine Nait-Ali	Maitre de conférences (HDR)	LISSI	Paris XII	Examineur

Remerciements

Je tiens tout d'abord à remercier très sincèrement Mme Catherine MARQUE, Professeur à l'Université de Technologie de Compiègne, Directrice de l'UMR CNRS 6600, pour avoir dirigé ce travail de recherche, pour m'avoir intégré dans son équipe de travail et m'avoir fait confiance tout au long de ces années.

Mes remerciements et ma sincère reconnaissance s'adressent à mon Co-directeur de thèse, Mr. Mohamad KHALIL, Chef du département de Génie Biomédical à l'Université Islamique du Liban pour m'avoir formé scientifiquement tout au long de ce travail de thèse. Je le remercie pour avoir crû en moi depuis le début de ma carrière scientifique. Je n'oublierai pas son enthousiasme, sa gaîté et son soutien moral dans les moments difficiles.

En témoignage de mon profond respect, je remercie Mr. Dimitri Lefebvre, Professeur à l'Université du Havre, et Mr. Guy Carrault, Professeur à l'Université de Rennes 1, de m'avoir fait l'honneur d'accepter de juger ce travail en tant que rapporteur.

Je remercie très vivement Mr Thierry DENOEU, Professeur à l'Université de Technologie de Compiègne, et Mr. Amine Nait-ali, Maître de conférences à l'Université Paris XII, qui ont accepté d'être les examinateurs de cette thèse.

J'exprimerai un remerciement tout particulier à Mme Chantal PEROT, la responsable de la formation de recherche en Génie Biomédical pour avoir veillé à l'amélioration du niveau des études et des travaux de recherche. Je n'oublierai pas toute l'équipe du laboratoire de Biomécanique et Génie Biomédical de l'UTC.

Je remercie tout spécialement le Président de l'Université Islamique du Liban (UIL), Mr Hassan CHALABI, Le Secrétaire Générale de l'UIL, Mr. Samih FAYAD, le Directeur de Faculté de génie de l'UIL, Mr. Abdel-Ménam KOUBAISSI, pour leurs soutiens, et pour leurs conseils.

J'ai effectué une partie de mon travail de thèse au sein du laboratoire Signaux, Images et Instrumentation (S2I) de l'Université Islamique du Liban où Je remercie mes collègues thésard Oussama MOUSTAPHA, Lina EL-KANSA, Mohamad AYACH, Walid FAHES et Jamal HAYDAR.

Je tiens à témoigner ma reconnaissance et mes sentiments les plus intimes à Mr. Gilbert FARGES, Enseignant- chercheur à l'UTC, pour les qualités humaines dont il fait preuve, ses critiques, ses remarques et ses conseils fort judicieux m'ont été profitables.

J'adresse également mes remerciements à mes amis Maher SABBAH, Walid HASSAN, Wissam JRADI, Mohamad KASSAB, Firas EL-KADRI, Gina SAAD et Nisrine TEMRAZ, sans qui tout aurait été très différent.

Ma gratitude s'étend à mes parents, mes beaux-parents, mon frère Ahmad, mes sœurs Hanadi Malak et Manar, mes beaux frères Niaz et Dali, ma belle sœur Nibras, son mari Mohamed, Saly et Nermine, à qui je dédie ce mémoire ainsi qu'à tous ceux qui partagent le sens de l'amour, la paix, la dévotion et le sacrifice.

Mes dernières pensées iront à ma fiancée Nijez qui a eu le courage de me supporter et m'encourager tout au long de cette aventure. Sa patience et son attention envers moi ne sont que des preuves de son amour.

Enfin, je baisse ma tête tout modestement devant le Miséricordieux qui donne tout et à Lui tout revient.

Résumé :

L'accouchement prématuré reste la principale cause de mortalité et de morbidité néonatales. Le signal EMG utérin semble un vecteur potentiel d'indication du risque d'accouchement prématuré.

Dans la suite des travaux réalisés pour la détection, le traitement et la classification des événements dans le signal EMG Utérin, notre travail s'est orienté vers la classification des contractions à partir des signaux EMG utérins, afin de séparer les deux types d'accouchement : accouchement prématuré et accouchement à terme.

Les contractions utérines ont été manuellement segmentées à partir du signal EMG utérin. Puis chaque contraction est modélisée et des paramètres sont extraits avant de faire la classification. Cette modélisation est faite par ondelettes et par analyse de la densité spectrale de puissance de chaque contraction.

La classification est ensuite réalisée en utilisant 2 types de méthodes : tout d'abord une classification non supervisée, qui regroupe les contractions sans connaissance a priori des classes, permettant ensuite une interprétation des groupes en fonction des semaines d'aménorrhées et du terme d'accouchement. Dans ce contexte nous avons développé une méthode originale de classification non supervisée basée sur le test de Fisher combiné avec la méthode de k-moyenne (USCM, Unsupervised Statistical Classification Method).

L'autre type de classification est supervisée. Après avoir sélectionné d'une façon précise les femmes qui peuvent être utilisées pour l'apprentissage de notre méthode de classification, nous avons utilisé différentes méthodes supervisées de classification. Tout d'abord, nous avons testé des méthodes classiques (Réseaux de neurones, Parzen,...). Puis une méthode originale basée sur le réseau d'ondelettes a été développée pour cette classification, cette méthode ayant été précédemment utilisée pour la régression mais jamais pour la classification.

Nous avons été confrontés à un problème lié au faible nombre de d'éléments pour l'apprentissage. Nous avons donc aussi utilisé une méthode basée sur la modélisation autorégressive pour augmenter l'ensemble d'apprentissage.

En ce qui concerne les applications, et pour la séparation entre les signaux d'EMG (application clinique), nous avons utilisé deux approches. Dans la première approche, nous avons utilisé des contractions ayant le même nombre de SAR (Semaines d'Aménorrhée à l'Enregistrement) mais des SAA (semaines d'Aménorrhée à l'accouchement) différent (petite différence et grande différence). La deuxième approche est de classifier les événements acquis avec différentes (SAR) pour des femmes ayant le même SAA.

D'après les résultats obtenus, nous avons pu conclure que nous pouvons distinguer le terme d'accouchement des femmes enregistrés au mêmes termes de grossesse. Et nous avons pu également conclure que les contractions changent de caractéristiques en fonction du terme de grossesse. D'un point de vue clinique, le résultat important est que, pour un terme de grossesse donné à l'enregistrement, il est possible de distinguer une contraction normale et une contraction conduisant à un accouchement prématuré.

Mots Clés: EMG Utérin, Contraction, Accouchement Prématuré, Classification Statistique, Réseaux d'Ondelettes, Modèle Autorégressif (Modèle AR).

Abstract:

Premature birth remains the main cause of neonatal mortality and morbidity. Uterine EMG signal seems a potential vector for the indication of the risk of premature birth.

In the continuation of the work completed for detection, the treatment and the classification of the events in Uterine EMG signal, our work was dedicated to the classification of the EMG contractions, in order to part the two types of labor: premature deliveries and deliveries at term.

The uterine contractions were manually segmented from the uterine EMG signals. Then each contraction is modeled and the parameters are extracted before making the classification. This modeling is made by wavelet and the analysis of the power spectral density of each contraction. Classification is then carried out by using 2 types of methods: first a non-supervised classification method, which is used to group the contractions with no a priori knowledge of the classes, and then permits to make an interpretation of the groups according to the weeks of gestation and term of delivery. In this context we have developed an original method of non supervised classification based on the Fisher test combined with the K-mean method (USCM, Unsupervised Statistical Classification Method).

The other type of used classification is supervised. After having selected in a precise way the women who can be used for the training of our method of classification, we used various supervised methods of classification. First, we used traditional methods (neural Networks, Parzen...). Then an original method based on the Wavelet network has been developed for this classification. This method had been previously used for the regression but never for classification.

We were faced with the low number of the set of training. We thus also developed a method based on autoregressive model (AR Model) to increase the training set.

Concerning the applications, and for separation between the EMG signals (clinical application), we used two approaches. In the first approach, we used contractions having the same RWG (Registration Week of Gestation) but different BWG (Birth Week of Gestation), by testing small and large differences. The second approach is to classify the events acquired with different RWG for women having the same BWG.

According to the results obtained, we can conclude that we can distinguish different delivery terms from recordings obtained from women having the same term of pregnancy. We can also conclude that the contractions change their characteristics according to the term of pregnancy.

In a clinical point of view, the important result is that we could distinguish, for a given recording term, normal contraction to contractions leading to the premature delivery.

Keywords: Uterine EMG, Contraction, Preterm deliveries, Statistical Classification, Wavelet Network (Wavnet), Autoregressive Model (AR Model).

Table des matières

Table des matières.....	7
Introduction générale.....	10
1.1 Introduction.....	12
1.2 La physiologie utérine.....	12
1.2.1 Muscle utérin.....	12
1.2.2 Activité utérine.....	14
Activité mécanique.....	14
Facteurs influençant l'activité utérine.....	14
1.3 Menaces d'accouchements prématurés MAP.....	15
1.4 Méthodes de Détection des MAP.....	16
1.4.1 Mesure de marqueurs biochimiques.....	16
1.4.2 Signes cliniques.....	17
1.4.3 Rôle des infections.....	17
1.4.4 Examen cervical et test de stimulation mammaire.....	18
1.4.5 Détection du rythme circadien de l'activité utérine.....	18
1.4.6 Tocographie externe.....	18
1.5 L'électromyographie utérine.....	21
1.5.1 Origine des signaux physiologiques.....	22
Excitation.....	22
Propagation.....	23
1.5.2 Acquisition et décomposition physiologique de l'EMG utérin.....	23
1.5.3 Caractéristiques détaillées des événements contenus dans l'EMG utérin.....	25
Contractions.....	25
Mouvements actifs fœtaux (MAF).....	27
Ondes d'Alvarez.....	27
Autres types.....	28
1.6 Traitement du signal EMG utérin.....	29
1.7 Détection des MAP par classification des EMG utérins.....	30
1.8 But de cette thèse.....	33
1.9 Conclusion.....	33
2.1 Introduction.....	34
2.2 La transformée en ondelette.....	34
2.2.1 Transformée en ondelettes continues.....	35
2.2.2 Transformée en ondelettes discrètes.....	36
2.2.3 Analyse multirésolution.....	37
2.2.4 Filtres de l'ondelette et de la fonction d'échelle.....	39
2.2.5 Propriétés et caractéristiques des ondelettes orthogonales.....	40
2.2.6 Extraction des paramètres.....	41
2.3 Modèle AR.....	41
2.3.1 Les méthodes d'estimation des paramètres AR.....	42
Méthode de Yule-Walker.....	42
Méthode de Burg.....	42
Méthode de Covariance.....	44
Méthode de Covariance modifiée.....	45
Problème de récursivité.....	45
2.3.2 Utilisation de la modélisation AR.....	46
2.4 Estimation spectrale pour l'EMG.....	47
2.4.1 Estimation non paramétrique.....	47
2.4.2 Estimation paramétrique.....	48
2.4.3 Estimation pseudo- spectrale par la méthode MUSIC.....	49
2.4.4 Discussion sur le choix de la méthode d'estimation du spectre des contractions.....	50
2.4.5 Extraction des paramètres à partir de la densité spectrale.....	51
Puissance de signal.....	51
Fréquence moyenne.....	51
Coefficient de dissymétrie.....	51
Coefficient d'aplatissement (Kurtosis).....	52
Fréquence médiane.....	52
Energie relative par bande de fréquence / Déciles.....	52
Fréquence pic.....	52
Entropie spectrale.....	53
Rapport H/B (haut/bas).....	53
2.5. Conclusion.....	54
3.1 Introduction.....	55
3.2 Classification supervisée.....	55
3.2.1 Méthode paramétrique : rapport de vraisemblance.....	56
.....	56

3.2.2 Méthode non paramétrique : Règle du K plus proche voisin.....	57
3.2.3 Méthode non paramétrique : Parzen.....	57
3.2.4 Les réseaux de neurones.....	58
Introduction.....	58
Réseaux de Neurones : Perceptron multicouches.....	59
Les réseaux à fonctions radiales (RBF pour Radial Basis Functions).....	65
3.3 Classification non supervisée.....	67
3.3.1 Méthode de K-moyenne.....	67
3.3.2 Méthode de Linde-Buzo-Gray (LBG).....	67
3.3.3 Méthode de K-moyenne flou.....	68
3.3.4 Réseaux de neurones compétitifs.....	68
Apprentissage compétitif.....	70
3.3.5 Classification statistique basé sur la loi de Fisher-Snedecor et test de comparaison de variances.....	70
Tests de Fisher-Snedecor pour échantillons indépendants.....	71
3.4 Méthode originale : méthode de classification non supervisée basée sur le test de Fisher.....	72
3.5 Analyse en composantes principales (ACP).....	73
3.6 Matrice de confusion.....	75
3.7 Conclusion.....	75
4.1 Introduction.....	77
4.2 Neurone formel à fonction dorsale.....	77
4.3 Définition des ondelettes multidimensionnelles et des réseaux d'ondelettes.....	78
4.3.1 Ondelettes multidimensionnelles.....	78
4.3.2 Réseaux d'ondelettes.....	78
4.3.3 Comparaison entre Réseaux d'ondelettes et réseaux de neurones.....	82
4.4 Apprentissage des réseaux d'ondelettes non bouclés.....	83
4.4.1 Calcul du gradient de la fonction de coût.....	83
4.4.2 Initialisation des paramètres du réseau.....	84
4.5. Réseaux issus de la transformée en ondelettes discrètes.....	86
4.5.1 Introduction.....	86
4.5.2 Structures obliques et bases d'ondelettes orthonormales.....	87
Ondelettes à variables continues.....	87
Ondelettes à variables discrètes.....	89
Choix de l'ondelette mère.....	89
4.5.3 Réseaux fondés sur la transformée discrète.....	90
4.6 Techniques de construction de réseaux d'ondelettes discrètes.....	90
4.6.1 Impossibilité d'utiliser les techniques de gradient.....	90
4.6.2 Approches utilisant une procédure de sélection.....	91
Technique fondée sur la construction de structures obliques étroites.....	91
4.7 Proposition d'une procédure de construction de réseaux et d'initialisation de l'apprentissage.....	91
4.7.1 Description de la procédure de construction de la bibliothèque.....	92
Famille engendrant la bibliothèque pour un modèle à une entrée.....	92
Cas des bibliothèques pour modèles à plusieurs entrées.....	93
4.7.2 Méthode de sélection par orthogonalisation.....	94
4.8 Conclusion.....	97
5.1 Introduction.....	98
5.2 Acquisition des signaux EMG utérins.....	98
5.3 Population.....	99
5.4 Classification Non supervisée.....	99
5.4.1 Méthodes choisies.....	99
5.4.2 Paramétrage à partir de décomposition par ondelettes discrètes.....	100
Extraction des paramètres.....	100
Choix de l'ondelette.....	100
Choix des échelles.....	100
5.4.3 Applications aux signaux synthétiques.....	101
Génération des signaux synthétiques.....	101
Résultats sur les signaux synthétiques.....	102
5.4.4 Application aux signaux réels.....	104
Résultats sur les groupes ayant même SAR et différents SAA.....	104
Résultats sur les groupes ayant même SAA et différents SAR.....	106
5.4.5 Conclusion sur les méthodes non supervisées.....	108
5.5 Classification Supervisée utilisant le Réseau à Fonctions Radiales et le Réseau d' Ondelettes.....	109
5.5.1 Méthodes choisies.....	109
5.5.2- Paramétrage à partir de la densité spectrale de puissance.....	109
5.5.3 Problème d'apprentissage et augmentation de la base de données.....	110
Choix de la méthode d'estimation du modèle AR.....	110
Choix de l'ordre du modèle et longueur bruit blanc gaussien.....	110
Génération des modèles APs et ATs.....	112
Réseaux d'ondelettes (Wavnet).....	113
Réseaux à fonction radiales (RBF).....	113
5.5.6 Application aux signaux réels.....	113
Classification utilisant les déciles de la DSP.....	113
Résultats sur les groupes ayant même SAR et différents SAA.....	113

Résultats sur les groupes ayant même SAA et différents SAR	114
5.6 Discussion.....	115
Conclusion générale.....	118
Référence.....	120
Publications relatives à ce travail.....	130

Introduction générale

L'accouchement prématuré demeure un problème de santé publique par ses conséquences médicales, psychologiques sur la mère (séparation mère/enfant), et par son retentissement économique [sena99]. Il a fait l'objet de nombreuses études depuis les années 1950 [dill46], [stee50], [hon58], [sure56], [val79], [wolf79], [plan84]. Malgré les progrès en obstétrique et en néonatalogie, l'accouchement prématuré constitue la plus grande menace pour le fœtus in utero. Il reste la principale cause de mortalité et de morbidité néonatales.

Le signal EMG utérin semble un vecteur potentiel d'indication du risque d'accouchement prématuré [lema99], [garf98]. Les perspectives de prédiction de la menace d'accouchement prématuré à partir de l'EMG utérin ont mis de nombreuses années d'études avant de se concrétiser.

Dans la suite des travaux réalisés pour la détection, le traitement et la classification des événements dans le signal EMG Utérin, notre travail s'est orienté vers la classification des contractions à partir des signaux EMG utérins, afin de séparer les deux types d'accouchement : accouchement prématuré et accouchement à terme.

Les bouffées d'activité électrique associées aux contractions utérines ont été manuellement segmentées à partir du signal EMG utérin. Puis chaque contraction est modélisée et des paramètres sont extraits avant de faire la classification. Cette modélisation est faite par ondelettes et par analyse de la densité spectrale de puissance de chaque contraction.

La classification est ensuite réalisée utilisant 2 types de méthodes : tout d'abord une classification non supervisée, qui est utilisée pour grouper les contractions puis faire une interprétation des groupes en fonction des semaines d'aménorrhées et du terme d'accouchement.

Ces méthodes non supervisées sont basées soit sur les réseaux de neurones compétitifs soit sur le test de Fisher combiné avec la méthode de k-moyenne (USCM, Unsupervised Statistical Classification Method) méthode originale développée dans ce travail.

L'autre type de classification est supervisée. Après avoir sélectionné d'une façon précise les femmes qui peuvent être utilisées pour l'apprentissage de notre méthode de classification, nous avons utilisé différentes méthodes supervisées de classification. Tout d'abord, les méthodes classiques sont utilisées (Réseaux de neurones, Parzen,...). Puis une méthode originale basée sur le réseau d'ondelettes a été développée pour cette classification. Cette méthode avait été précédemment utilisée pour la régression mais jamais pour la classification.

Nous avons été confrontés au faible nombre de l'ensemble d'apprentissage. Nous avons donc aussi utilisé une méthode basée sur la modélisation autorégressive pour augmenter l'ensemble d'apprentissage.

Ce mémoire est organisé de la manière suivante :

Le chapitre 1 est une présentation globale du problème traité. Nous commençons par la description de l'objectif global de l'étude, puis nous abordons la physiologie associée aux signaux EMG utérins. Ensuite nous décrivons les menaces d'accouchements prématurés (MAP) et les différentes méthodes de détection de ces MAP. Nous présenterons ensuite la forme et le contenu des signaux EMG et quelques caractéristiques des événements contenus dans l'EMG utérin. La dernière partie de ce chapitre est dédiée à la présentation des différentes étapes mises en œuvre dans les études précédentes pour le traitement et la classification des contractions.

Le chapitre 2 est dédié à la modélisation des signaux et à l'extraction des paramètres. Nous commençons par la décomposition en ondelettes discrètes et l'analyse multirésolution. Puis nous décrivons les paramètres extraits à partir des coefficients d'ondelettes.

La section 2 de ce chapitre est dédiée à la modélisation AR (le choix de l'ordre de modèle), puis aux méthodes d'estimation de la densité spectrale de l'EMG. La représentation du contenu spectral reposant sur l'estimation par le périodogramme moyenné, des méthodes des signaux paramétriques sur une méthode pseudo- spectrale (par la méthode MUSIC) sont discutées dans cette section. A la fin de cette modélisation, les paramètres extraits seront présentés.

Le chapitre 3 donne une présentation des différentes méthodes de classification : méthodes de classification supervisés et non-supervisés. Nous commençons par la description des réseaux de neurones artificiels, les réseaux à fonctions radiales et le perceptron multicouches, puis ensuite celle de la méthode du K plus proche voisin et de la méthode de Parzen. La dernière partie de ce chapitre est dédiée à la présentation des différentes méthodes de classification non supervisés. Nous commençons par la description de la méthode K moyenne puis des réseaux de neurones compétitifs et de leurs apprentissages. Nous présentons ensuite une méthode originale de classification statistique basée sur la loi de Fisher appliqué après décomposition en ondelette. La technique d'analyse en composante principale est présentée et est utilisée pour projeter les données dans un espace des paramètres réduits (généralement R^2) pour faciliter l'interprétation et la visualisation des résultats.

La nouvelle idée utilisée dans ce travail est présentée au chapitre 4. Plutôt que d'utiliser une décomposition en ondelettes, suivie par une classification par un réseau de neurones, nous avons directement utilisé un réseau d'ondelette, appliqué aux paramètres de la DSP de l'EMG. Il s'agit d'un réseau de neurones ayant des fonctions d'activation qui ressemblent aux ondelettes (dilatations et translations). Ce réseau a jusqu'alors été largement utilisé pour la régression. L'originalité est ici de l'adapter pour effectuer la classification.

Le chapitre 5 est consacré à la présentation des résultats après avoir appliqué les différents algorithmes et méthodes développés sur des signaux simulés et des signaux réels enregistrés au cours de la grossesse.

Après validation des méthodes sur les signaux synthétiques, les résultats de classification sur les signaux réels sont présentés en fonction du terme à l'enregistrement et du terme à l'accouchement : classification sur les groupes possédant les mêmes Semaines d'Aménorrhées à l'Enregistrements (SAR) et différents Semaines d'Aménorrhées à l'Accouchement (SAA), puis classification sur des groupes possèdent les mêmes SAA et différents SAR.

Les méthodes de classification supervisées et non-supervisées sont appliquées ensuite sur ces groupes et les résultats sont montrés dans la troisième section de ce chapitre. Une interprétation des différents résultats conclut ce chapitre.

Ce mémoire se termine sur une discussion et une conclusion générales, amenant aux perspectives de ce travail.

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 classification des contractions dans le signal électromyographique (EMG) utérin, signal présentant des caractéristiques non stationnaires. L'objectif final de ce travail étant la prévention de l'accouchement prématuré.

Dans ce chapitre, nous commençons par la description de l'objectif global de l'étude, puis nous abordons brièvement la physiologie associée aux signaux EMG utérin. Nous parlerons ensuite des menaces d'accouchements prématurés (MAP) et des différentes méthodes de détection de MAP. Nous présenterons ensuite la forme et le contenu des signaux EMG et quelques caractéristiques des événements contenus dans l'EMG utérin.

La dernière partie de ce chapitre est dédiée à la présentation des différentes étapes de traitement qui ont été jusqu'alors mises en œuvre pour la classification de l'EMG utérin.

1.2 La physiologie utérine

1.2.1 Muscle utérin

L'utérus est un organe creux en forme de poire renversée. Avant la première grossesse, l'utérus adulte mesure environ 7,5 cm de longueur, 5 cm de largeur et 2,5 d'épaisseur.

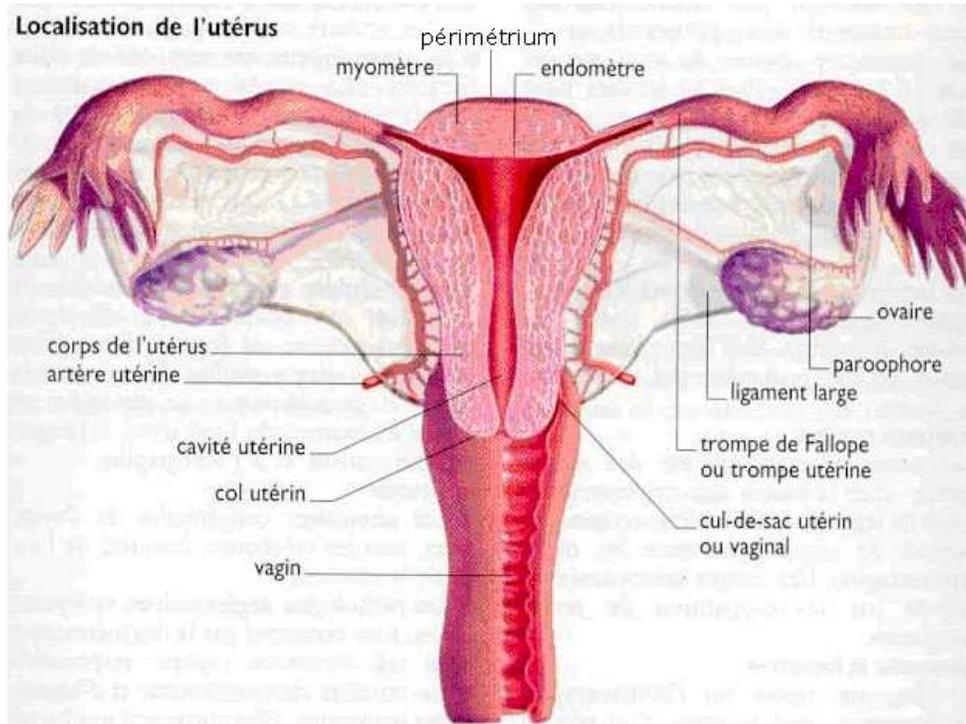
Les divisions anatomiques de l'utérus sont le fond, la partie en forme de dôme située au dessus des trompes de Fallope, le corps, la partie principale, centrale et aplatie, et le col, l'étroite partie inférieure qui s'ouvre sur le vagin. Entre le corps et le col se trouve l'isthme, une région comprimée d'environ 1 cm de longueur.

Normalement, l'utérus est fléchi entre le corps et le col dans une position d'antéversion. Dans cette position, le corps de l'utérus se projette vers l'avant et légèrement vers le haut au dessus de la vessie, et le col se projette vers le bas et vers l'arrière et entre dans la paroi antérieure du vagin à un angle presque droit. Plusieurs ligaments qui sont des prolongements soit du péritoine pariétal, soit des cordons fibro-musculaires, maintiennent la position de l'utérus.

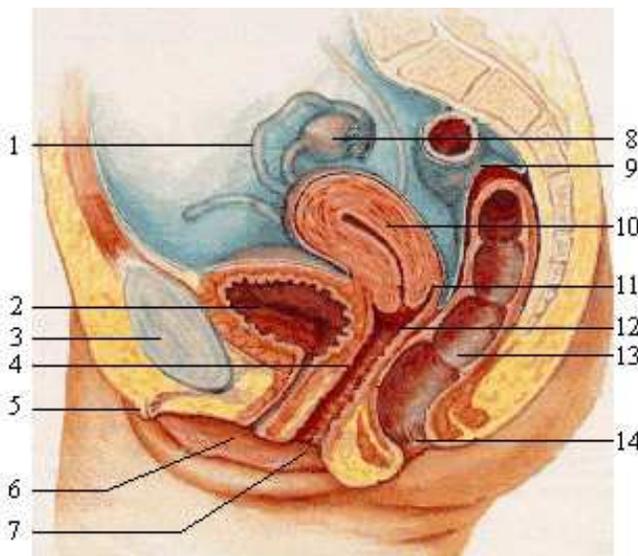
Bien que les ligaments maintiennent normalement la position d'antéversion de l'utérus, ils laissent également au corps utérin une certaine liberté de mouvement. Par conséquent, un déplacement de l'utérus peut se produire. Une rétroversion se produit lorsque l'utérus bascule vers l'arrière.

Sur le plan histologique, l'utérus est composé de trois couches de tissus [tort94] : le périmétrium, le myomètre et l'endomètre. La couche externe, le **périmétrium** ou séreuse, fait partie du péritoine viscéral (figure 1.1). Sur le côté, il devient ligament large. Vers l'avant, il passe au-dessus de la vessie et forme une poche peu profonde, le cul-de-sac

vésico-utérin. A l'arrière, il passe au-dessus du rectum et forme le cul-de-sac recto-utérin (de Douglas), le point le plus bas de la cavité pelvienne.



Vue postérieure



1. trompes de Fallope
2. vessie urinaire
3. symphyse pubienne
4. ligaments vésico-pubiens
5. clitoris
6. petite lèvre
7. vestibule vaginal
8. ovaire
9. péritoine
10. utérus
11. cul-de-sac vaginal
12. col
13. rectum
14. anus

Figure 1.1. Structure de l'utérus

La couche intermédiaire de l'utérus, le myomètre, forme la plus grande partie de la paroi utérine. Composé de trois couches de fibres musculaires lisses, le myomètre est plus épais au niveau du fond et plus mince au niveau du col. Au cours de l'accouchement, en réponse à différents facteurs dont l'ocytocine qui provient de la neurohypophyse, les contractions coordonnées des muscles aident à expulser le fœtus du corps de l'utérus.

La couche interne, l'endomètre, est richement vascularisée et composée d'une couche superficielle d'épithélium cylindrique simple (composée de cellules ciliées et sécrétrices), de glandes utérines (endométrales), qui sont des invaginations de l'épithélium de surface et du stroma, une région très épaisse de chorion (tissu conjonctif).

1.2.2 Activité utérine

Activité mécanique

L'activité mécanique de l'utérus non gravide est cyclique et dépend du cycle hormonal menstruel. Elle comprend une phase d'activité propagée dans tout l'utérus qui permet l'expulsion du sang et des débris de l'endomètre. Le début du cycle est donc caractérisé par des contractions propagées à tout l'utérus, dirigées vers le col et de fréquence 1-3/min. Le milieu du cycle, période ovulaire, est caractérisé par de fortes contractions régulières dirigées vers le fond, de fréquence 10/min, et dont on pense qu'elles favorisent la progression du sperme. Suit une phase quiescente qui correspond au développement de l'endomètre nécessaire à une éventuelle nidation, les contractions sont locales, de faible intensité, dirigées vers le fond et de fréquence 265/min [chall93].

L'utérus gravide comprend lui aussi une phase de relative quiescence pendant la majeure partie de la grossesse, suivie par une période d'activité aboutissant à l'accouchement.

Chez la femme, deux types d'activités utérines cohabitent :

- des contractions de très faible amplitude, de fréquence 1/min et d'influence très locale, nommées Onde Alvarez (OA).

- des contractions de plus forte amplitude et de plus faible fréquence (une toutes les 3 ou 4 heures au début de leur apparition vers 18 semaines d'aménorrhée). Ces contractions, appelées contractions de Braxton Hicks deviennent de plus en plus fortes et fréquentes à l'approche du terme (1 par heure vers la 30^{ième} semaine) et ont un champ d'influence plus étendu que les ondes d'Alvarez [wood72], [merg79].

Parallèlement à l'évolution de ces contractions, les phénomènes de maturation cervicale débutent puisque le col commence à se ramollir et à s'effacer.

La modification la plus importante notée en début de travail concerne la propagation de l'activité. De partiellement propagées en fin de grossesse, les contractions de travail deviennent fortes, rythmées et sont propagées à tout l'utérus. Dans un premier temps, ces contractions ont pour effet de raccourcir le col, pour ensuite l'effacer et le dilater totalement. En même temps les contractions poussent le fœtus jusqu'à permettre son expulsion.

Facteurs influençant l'activité utérine

L'activité contractile de l'utérus est principalement dominée par l'influence d'hormones dont les actions sont explicitées ci-dessous :

- stéroïdes ovariens (œstrogène et progestérone) : au début de la grossesse, les œstrogènes et la progestérone sont sécrétés au niveau ovarien par le corps jaune formé après l'ovulation. Dans les 6 derniers mois de la grossesse, la production de ces stéroïdes est entièrement due à la sécrétion placentaire. Juste avant le travail, on observe une augmentation du rapport œstrogène/progestérone [fuch73]. Cependant, le niveau urinaire ou plasmatique de ces hormones ne représente pas directement leur concentration au niveau myométrial [parl84].

Les œstrogènes ont pour effet de provoquer une hypertrophie utérine et d'augmenter la concentration en protéines contractiles. Ils favorisent aussi la formation de « gap junctions » [garf80]. Ces « gap junctions » sont des jonctions permettant la communication

intracellulaire directe. Leur rôle principal est donc de faciliter la propagation de l'activité contractile.

- l'ocytocine est produite par la glande pituitaire. Dans l'espèce humaine, la sensibilité à l'ocytocine augmente progressivement au cours de la grossesse, jusqu'à atteindre un pic peu de temps avant l'accouchement. Cette augmentation est corrélée à l'augmentation des récepteurs à l'ocytocine dans le myomètre [fuch73].

L'ocytocine est fréquemment utilisée, soit pour provoquer un déclenchement artificiel du travail, soit pour accélérer le déroulement d'un accouchement normal. La réponse du myomètre à l'administration d'ocytocine exogène est fonction de la dose administrée, des débits trop élevés provoquant notamment une hyperactivité utérine [wolf79].

- il semblerait que la sensibilité de l'utérus aux prostaglandines PGE₂ et PGF_{2α} augmente pendant la grossesse. L'activation du travail semble correspondre à une augmentation du taux de prostaglandines, bien que l'on n'ait pas pu mettre en évidence de pics de PGE₂ et PGF_{2α} dans le plasma avant l'accouchement. Les taux de PGE₂ et PGF_{2α} mesurés dans le liquide amniotique augmenteraient pendant le travail. Cette augmentation serait sous le contrôle de la sécrétion d'ocytocine [fuch73], [ligg79].

Les prostaglandines sont de puissants utérostimulants. Elles sont largement utilisées pour provoquer des interruptions de grossesse avant terme ainsi que des accouchements à terme sur des cols non « mûrs », du fait de leur double action sur la contractilité utérine et sur la maturation cervicale [ligg79], [cald80], [audr84].

1.3 Menaces d'accouchements prématurés MAP

Le taux de prématurité (< 37 semaines d'aménorrhée (SA)) a peu évolué au cours des 20 dernières années. Seules la France [brea95], la Finlande et la Norvège [olse95] ont observé une diminution entre la fin des années 60 et les années 80. En France, le taux de prématurité s'est stabilisé entre 1990 et 1995 ; il était de 5,9 % en 1995 [blon97], avant d'augmenter dernièrement (6,8 % en 1998) [blon01]. Il atteint 7 % au Canada et plus de 10 % aux États-Unis [heam01].

La France, pays industrialisé de la communauté européenne, compte 58,7 millions d'habitants dont 51,3% de femmes. Le nombre moyen d'enfants par femme était de 1,75 en 1998 (Moyenne européenne = 1,44 enfants/femme), ce qui correspond à 7 040 300 nouveaux nés par an [kerj99] dont 5,9% de prématurés (âge gestationnel inférieur à 37 semaines d'aménorrhée) et 1,6% de grands prématurés (âge gestationnel inférieur à 34 semaines d'aménorrhée) [dree98]. Grâce à un suivi important des femmes pendant leur grossesse, associé au développement de la surveillance périnatale, la mortalité des nouveaux nés au cours de leur premières semaines ou de leur premier mois d'existence est pratiquement stabilisée à un niveau plancher depuis 1990. Avec 13 décès pour 100 000 naissances, la France reste cependant la lanterne rouge de l'Europe occidentale. La raison en est, entre autre, que 84% des femmes à risque [sesi98] accouchent encore dans des maternités dépourvues d'un service de réanimation néonatale, et que le transfert postnatal par SMUR (concerne 8,7% des naissances [dree98]) se révèle plus dangereux que celui de la femme enceinte. Il est donc nécessaire de repérer au plus tôt les grossesses à risque pour les orienter vers les maternités les mieux équipées [rous98]. La question est donc comment mieux développer la technique de prévention de la menace d'accouchement prématuré (MAP).

La menace d'accouchement prématuré se situe au premier rang des complications de la grossesse. Elle est une des principales causes d'hospitalisation au cours de la grossesse. En France en 1998, près de 20 % des femmes ont été hospitalisées au cours de leur grossesse. Parmi elles, près de 38 % l'ont été en raison d'une menace d'accouchement prématuré [blon00]. Haas et al. rapportent des chiffres comparables aux États-Unis [haas96]. Parmi les

femmes hospitalisées au cours de la grossesse (18 %), le travail prématuré était la première cause d'hospitalisation (29 %), loin devant l'hypertension (13 %), le diabète (10 %), les hémorragies (9 %) et la rupture prématurée des membranes (7 %). La menace d'accouchement prématuré représente aussi une situation conduisant à l'accouchement prématuré. Sur une population de 239 femmes présentant une menace d'accouchement prématuré, Hueston *et al.* [hues98] montrent que 33 % ont accouché prématurément (dont 67 % après 33 SA). Bien que les études d'observation sur le risque d'accouchement prématuré chez des femmes présentant une MAP soient peu nombreuses, des informations issues des essais cliniques sont disponibles. Cela concerne les travaux sur les nouveaux marqueurs de risque de prématurité, en particulier la fibronectine. Ainsi, dans des populations symptomatiques, c'est à dire présentant des contractions utérines et des modifications cliniques du col, le risque d'accouchement avant 37 SA varie de 20 % à 67 % selon les études [goff01]. Les essais thérapeutiques randomisés sur l'efficacité des tocolytiques montrent que 50 % des femmes du groupe placebo accouchent à terme [king88]. Dans ces essais, les critères diagnostiques sont standardisés. Dans la pratique clinique courante, où les modalités du diagnostic de MAP sont moins strictes, la proportion de femmes diagnostiquées « MAP » accouchant à terme est probablement plus élevée. Quoi qu'il en soit, ces résultats donnent la mesure des effets sur la santé et de l'impact en termes de coûts, de la menace d'accouchement prématuré et de ses conséquences.

1.4 Méthodes de Détection des MAP

1.4.1 Mesure de marqueurs biochimiques

Le travail prématuré comporte divers étiologies et probablement différents mécanismes pathologiques. Ces différents mécanismes pourraient aboutir néanmoins à une voie pathologique commune conduisant à l'initiation du travail. Certaines équipes ont supposé que ceci pourrait être reconnu grâce à des marqueurs biochimiques.

D'après l'étude de Darne en 1987 [darn 87], il semble qu'il y ait une relation entre l'augmentation du rapport estriol/progestérone (concentration salivaire) et la baisse du rapport estradiol/progestérone pour les femmes en travail prématuré. Cependant, ce test est utilisé au plus 4 jours avant l'accouchement et ne peut offrir une opportunité de prédiction du travail prématuré à un stade précoce.

Dans les cas de travail prématuré associé à une infection intra-utérine, l'élévation dans le liquide amniotique ou le plasma de substances telles les cytokines, phospholipase A2, pourrait prédire la survenue d'un travail prématuré [rome89], [hall89] et [bren90]. Néanmoins, ces médiateurs de l'inflammation ne semblent pas avoir d'intérêt lorsque le travail n'est pas en rapport avec une infection. Ils ne permettent pas non plus d'identifier les femmes à haut risque.

De même, certains peptides placentaires (CRH corticotropin releasing hormon) sont présents en concentration importante plusieurs semaines avant le début du travail dans le plasma maternel pendant ou avant le travail prématuré [camp87]. Il s'agit néanmoins d'études préliminaires qui doivent être validées sur de plus grandes séries.

Parmi les nouveaux marqueurs, la fibronectine fœtale possède une sensibilité faible (21%) en population générale, mais sa spécificité est correcte (98%) [peac97].

Ces marqueurs ont probablement une valeur diagnostique, mais ils n'ont pas encore permis de diminuer la prématurité dans la population générale. De plus, leur stratégie d'utilisation n'est pas encore bien définie [beno99].

1.4.2 Signes cliniques

Katz en 1990 [katz90] avait interrogé 200 patientes sur la présence, la durée et les caractéristiques de plusieurs symptômes. Dans les 7 jours précédant le travail prématuré, les femmes qui ont accouché prématurément ont ressenti significativement plus de contractions, de crampes de type menstruel, de douleurs dorsales, de pression pelvienne, de pertes vaginales, que les femmes témoins. Pour Copper en 1990 [copp90], les signes de diarrhée sont les plus caractéristiques de l'AP.

Ces signes pourraient, bien entendu, constituer une aide au diagnostic précoce du travail prématuré sans que cela puisse constituer une forme de surveillance unique. En effet, dans près de 30% des cas, le travail prématuré survient sans qu'aucun signe clinique particulier n'ait pu être noté.

Ces signes ou symptômes, de par leur caractère suggestif et leur valeur prédictive positive ou négative assez faible, ne peuvent constituer la base d'un programme de prévention de la menace d'accouchement prématuré.

1.4.3 Rôle des infections

Romero et Al. ont montré en 1989 que 30% des accouchements prématurés sont associés à un processus infectieux [rome89]. Les pathologies en cause sont essentiellement les infections cervico-vaginales et urinaires. Mais toute autre infection systémique peut aussi être impliquée.

Trois à 5% des femmes enceintes développent une bactériurie asymptomatique en début de grossesse. Leur risque d'accoucher prématurément est multiplié par deux par rapport à la population générale [beno99]. Seul 1/3 d'entre elles sont symptomatiques. Or, toute pyélonéphrite aiguë non traitée est associée dans 30% des cas à un AP. Le dépistage et le traitement des infections urinaires permettent donc de réduire le risque de prématurité.

Cinq à 30% des femmes enceintes sont porteuses asymptomatiques de streptocoques B, et le taux d'AP est de 20% dans cette population. La présence de ce germe est également associée à un risque élevé de rupture prématurée des membranes avant terme. L'antibiothérapie en cas de MAP associée à une infection à streptocoque B s'est montrée efficace pour réduire le taux d'AP.

Le mycoplasme est fréquemment retrouvé au niveau cervico-vaginal. Il peut être responsable de chorio-amniotites, d'hypotrophie, d'avortements spontanés à répétition et de mort fœtale in utero. Mais son rôle direct dans la prématurité reste discuté.

L'isolement de *Chlamydia trachomatis* dans un prélèvement vaginal est lié à une augmentation des risques d'avortements spontanés, d'hypotrophie et d'accouchement prématuré. L'efficacité du traitement antibiotique est controversée.

Enfin, plusieurs études ont retrouvé une association entre vaginose bactérienne et travail prématuré, mais les résultats ne sont pas toujours significatifs [beno99].

1.4.4 Examen cervical et test de stimulation mammaire

Depuis plusieurs années, la dilatation cervicale prématurée et l'effacement du col sont associés à un haut risque d'accouchement prématuré. Or, la littérature est confuse à ce sujet. Certains auteurs ont trouvé une association entre les changements cervicaux et l'AP [papi86], alors que d'autres auteurs ont conclu que la dilatation cervicale au début du 3^{ème} trimestre est une variante anatomique normale spécialement chez la multipare [floy61], [pari61]. Dans une étude postérieure, Leveno et Al. en 1986 [leve86], ont montré que 4 des 15 femmes étudiées (27%), ayant une dilatation cervicale de 2 cm ou plus dans le troisième trimestre de la grossesse avaient accouché prématurément.

Guin et Al. en 1994 [guin94] ont montré que la combinaison d'un test de stimulation mammaire et de l'examen cervical a permis d'identifier 35% des accouchements prématurés chez les patientes nullipares. Ce test avait été effectué sur 267 patientes nullipares entre 26 et 28 semaines.

La longueur du col peut aussi être étudiée par ultrasonographie endovaginale. Elle est normale de 30 à 50 mm Jusqu'à 24 SA, et de 26 à 40 mm après 24 SA. C'est une technique sensible et reproductible mais elle est chère et invasive et donc peu utilisable comme test de routine [acog88], [mcle93], [dupu99].

1.4.5 Détection du rythme circadien de l'activité utérine

D'après l'étude de Germain en 1993 [germ93], chez les femmes ayant accouché à terme, il existe à partir de 80 jours avant l'accouchement un pic nocturne de l'activité utérine. La disparition de ce pic nocturne à partir de 24 jours avant l'accouchement est liée au travail prématuré.

Il s'agit d'une étude préliminaire sur un petit groupe de patientes qui constitue une piste de recherche très intéressante.

1.4.6 Tocographie externe

Actuellement, le seul outil permettant de mesurer l'activité utérine est l'enregistrement externe mécanique, la tocographie externe. Cet appareil permet l'enregistrement de l'activité mécanique de l'utérus par mesure de la déformation de l'abdomen sous l'effet d'une contraction. De nombreuses études se sont attachées à décrire l'activité utérine pendant la grossesse par tocographie externe. Nageotte a réalisée, en 1988, une étude sur 237 patientes à bas risque d'accouchement prématuré, dont les contractions ont été enregistrées à l'aide d'un tocographe ambulatoire [nage88]. Il a remarqué une progression significative de l'activité utérine de 4,7% par semaine, entre 30 et 44 semaine d'aménorrhée (SA), comme le montre la figure 1.2.

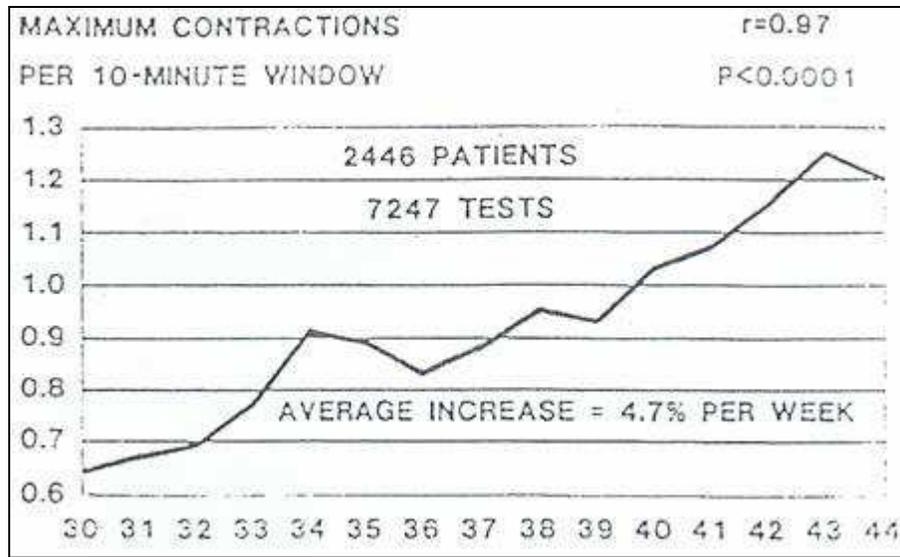


Figure 1.2. Augmentation de la fréquence des contractions selon le terme [nage88].

L'activité utérine précédant le travail était plus importante dans le groupe des accouchements prématurés (AP) que dans le groupe des accouchements à terme (AT) et moins importante dans le groupe des femmes ayant accouché à terme dépassé (ATD), comme l'indique la figure 1.3.

Trois jours avant le début du travail spontané, l'activité utérine augmente brutalement (effet de crescendo) dans les trois groupes (figure 1.4).

Les conclusions sont identiques dans l'étude de Katz et Al. en 1986, qui ont enregistré 34 femmes à haut risque par tocographie ambulatoire [katz86]. Ils concluent en effet que la fréquence d'apparition des contractions est plus importante chez les femmes ayant accouché prématurément.

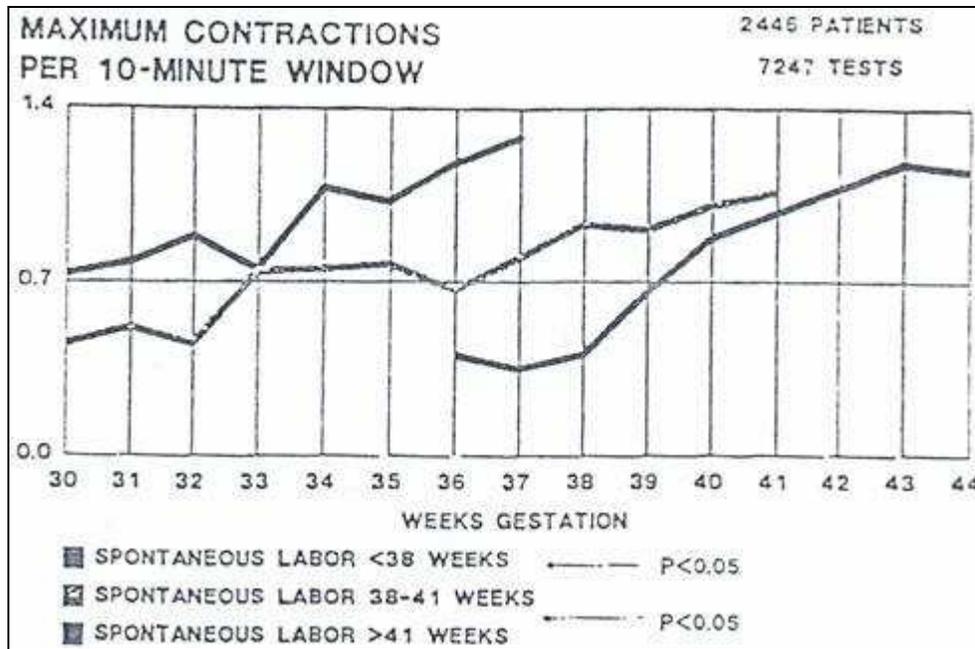


Figure 1.4. Activité utérin chez les patientes qui ont accouché prématurément, à terme, à terme dépassé [nage88].

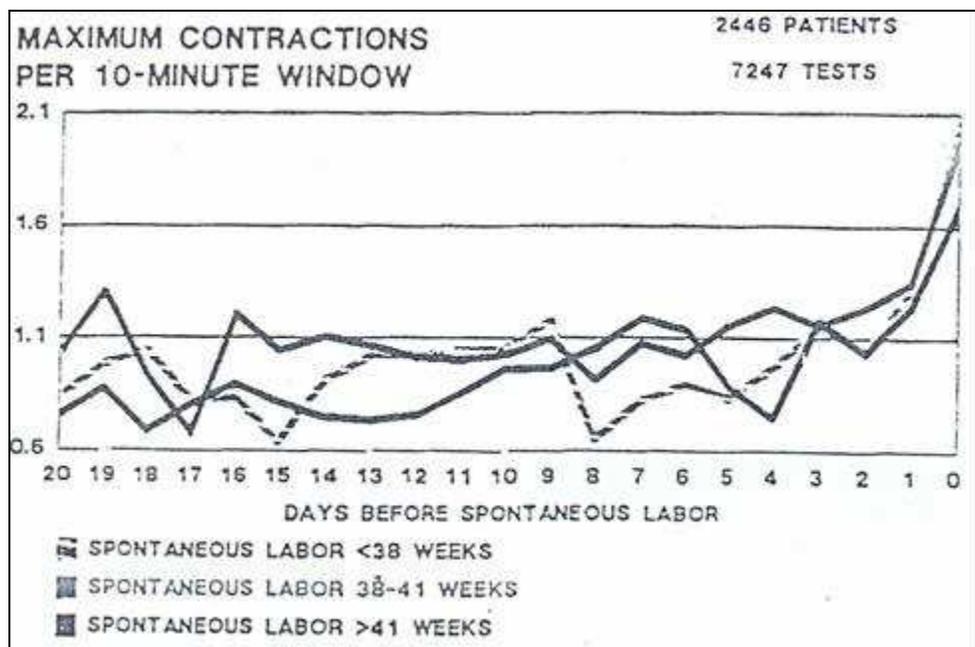


Figure 1.4. Effet de « crescendo » de l'activité utérine pendant les 3 jours précédant le travail [nage88].

Les comparaisons des différences dans la fréquence d'apparition des contractions est significative entre les trois groupes AP, AT et ATD.

La littérature est néanmoins confuse quand il s'agit de quantifier cette fréquence d'apparition. Saling en 1982 [sali82], Klock et Chantraine [kloc75], Koepke et al [koep75] en 1975, considèrent que 3 contractions par heure pendant la grossesse est un phénomène normal alors que Ruttgers et al en 1972 [rutt72], Fisher en 1973 [fish73], et d'autres auteurs considèrent plus de 3 contractions par heure comme un phénomène encore normal. Quant à Langreder en 1962 [lang62], il situe la normalité à 16 contractions par heure vers la fin de

la grossesse. Zahn en 1984 [zahn84] a fait une étude sur 11 femmes à bas risque d'accouchement prématuré, enregistrées pendant 24 heures par tocographie. A la 30^{ème} semaine, les femmes ressentait environs 2 contractions par heure. A partir de la 38^{ème} semaine, elles ressentait plus de 3 contractions par heure.

Il met aussi en évidence, comme Germain en 1993 [germ93], un pic nocturne de l'activité utérine entre 22h30 et 2h00 du matin et montre que c'est le matin que la fréquence d'apparition des contractions est minimum.

Zahn conclue que les résultats ne sont pas statistiquement significatifs, et que la fréquence d'apparition des contractions utérines n'est pas un signe clinique suffisant pour détecter les MAP.

Enfin, Newman en 1987, a réalisé une étude sur la signification des contractions appelées ondes d'Alvarez (OA) qui ont des propriétés particulières [newm87] : elles possèdent une fréquence d'apparition élevée et sont de faible amplitude. Elles sont connues pour être locales car ne se propageant pas à tout l'utérus.

Cent quarante deux femmes ont été enregistrées grâce à un tocodynamomètre ambulatoire entre 23 et 36 SA. Les patientes à risque d'accouchement prématuré ont présenté significativement plus d'OA que les femmes à bas risque. Ni la parité, ni l'âge gestationnel n'ont d'effet sur la fréquence d'apparition des OA. Par contre, les OA sont plus fréquentes en présence de gestation multifoetale, indépendamment de l'accouchement prématuré.

Il montre aussi que la thérapie tocolytique provoque une diminution de 50% du nombre d'OA.

L'association apparente entre ces OA et l'apparition du travail prématuré a mené certains investigateurs [crea84] à penser que ces ondes traduisent une irritabilité utérine qu'Alvarez qualifiait de fibrillation utérine. Bien que cette irritabilité puisse causer ou non une dilatation ou un effacement cervical, elle pourrait permettre d'identifier les femmes à risque. Il est peu probablement que les OA seules provoquent la dilatation ou l'effacement du col.

L'utilisation de la fréquence d'apparition des contractions, seul paramètre actuellement offert par les systèmes de surveillance de la grossesse (tocographie externe), semble donc inefficace pour une prévention valable des accouchements prématurés. Nous voulons donc, dans cette étude, travailler sur les caractéristiques de chaque contraction, afin de pouvoir discriminer de manière fiable les contractions normales de grossesse des contractions pathologiques. Ces caractéristiques seront extraites de la bouffée d'activité électrique, associée à chaque contraction, recueillie de manière non invasive sur l'abdomen de la femme (électrohystérogramme : EHG, ou électromyographie utérine : EMG utérin).

1.5 L'électromyographie utérine

L'étude de la contractilité utérine par électromyographie (EMG) a été entamée depuis de nombreuses années. L'une des premières études date de 1931 [bode31]. Les premières études se sont concentrées sur l'analyse temporelle puis fréquentielle de ce signal. Le signal EMG semble être un vecteur potentiel d'indication du risque d'accouchement prématuré. En effet, il existe une bonne corrélation entre la survenue de bouffée d'activité électrique sur ce signal et l'augmentation de la pression intra-utérine [stee50]. 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 méthode de traitements du signal EMG en vue de l'extraction de paramètres caractéristiques des contractions, de leurs efficacités,...

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 μ V et 1,8 mV [Lope84]. 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] [wolf70]. 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 [mans91].

L'EMG, enregistré à l'aide d'électrodes de surface positionnées sur l'abdomen, se caractérise par une activité basse fréquence (0,03 - 0,1 Hz), sur laquelle se superposent des bouffées d'activité de plus haute fréquence (0,3 - 2Hz). Le signal de basse fréquence est considéré comme une activité résultant des perturbations mécaniques induites pour la déformation du ventre sous l'effet de contractions [sure56].

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 [marq86]. 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'avancée du terme de la grossesse chez la femme [marq86]. 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.

1.5.1 Origine des signaux physiologiques

L'activité électrique étant à l'origine de l'effet mécanique, ce dernier dépend de deux paramètres régissant le processus contractile : l'excitation et la propagation de l'activité électrique.

Excitation

Etudiée au niveau cellulaire, l'activité électrique peut être caractérisée à l'aide de deux types de potentiels, le potentiel de repos et le potentiel d'action.

Potentiel de repos : c'est la différence de potentiel entre le potentiel intérieur négatif et le potentiel extérieur positif d'une cellule au repos, sachant que ce potentiel est directement lié aux concentrations des ions calcium, sodium, potassium et chlorure. Chez la femme enceinte, il est de l'ordre -65 à -80 mV [kuma61].

Potentiel d'action (PA) : le PA est dû à de brusques variations de la perméabilité des parois de la cellule aux ions précités, et correspond à une diminution de la positivité du potentiel extérieur à la cellule. Au delà d'un certain seuil, des PA sont générés. Lors des contractions utérines, ces PA sont regroupés par bouffées que l'on peut caractériser à l'aide de deux type de fréquences : la fréquence inter-bouffée (< 0.1 Hz) qui reflète l'excitabilité des cellules, alors que la fréquence intra-bouffée (de 1 à 10 Hz) conditionne l'intensité de la contraction mécanique résultante [deve93]. Or il est très intéressant de constater (comme indiqué sur la figure 1.5) que la fréquence intra-bouffée n'est pas constante au cours de la bouffée, puisqu'elle augmente en milieu de bouffée pour ensuite rediminuer.

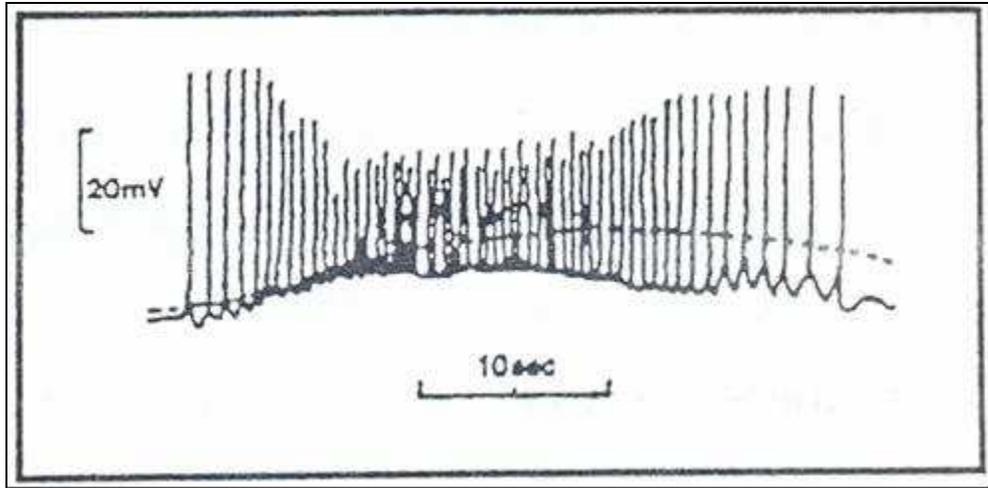


Figure 1.5. Bouffée de PA enregistrée sur un rat à l'aide d'une microélectrode [kuri61].

Propagation

Une cellule myométriale peut soit être excitée par les PA émanant d'une cellule voisine (on l'appelle alors 'pacefollower'), soit générer ses propres potentiels (c'est alors une 'pacemaker'). On peut reconnaître un potentiel généré par une cellule 'pacemaker' du fait de la lente dépolarisation, appelée prépotentiel, qui le précède. Ce prépotentiel est absent dans le cas d'une cellule 'pacefollower' [kuri61], [kao77]. Cependant, contrairement aux cellules cardiaques, les cellules myométriales peuvent alternativement être 'pacemaker' ou 'pacefollower'. De plus, il ne semble pas y avoir de zones 'pacemaker' fixes : une bouffée d'activité peut être générée par n'importe quelle cellule, sachant que les cellules 'pacemaker' peuvent changer d'une contraction à l'autre.

De plus, le couplage électrique inter-cellules est amélioré par la présence de 'gap junctions'. Ces 'gap junctions' sont des zones où les membranes de deux cellules adjacentes forment des pores permettant un couplage électrique entre elles. La mise en évidence des 'gap junctions' a été décrite par plusieurs auteurs [garf84], [sims82].

1.5.2 Acquisition et décomposition physiologique de l'EMG utérin

Les signaux électromyographiques observés peuvent être représentés par un processus aléatoire $x(t)$ de la forme [Khal99]:

$$x(t) = EMG(t) + \sum_{i=1}^n S_i(t) + n(t)$$

Cette relation modélise l'ensemble des activités significatives (signal utile, tracé de fond, bruit, artefacts...) qui constituent le signal. $EMG(t)$ est un modèle variable en fonction de l'état du muscle (repos, contraction) et en fonction du terme de l'analyse (évolution physiologique). Cette activité utérine de base $EMG(t)$ correspond à l'intégration filtrée de nombreuses activités électriques cellulaires. Chacune de ces activités élémentaires peut être considérée comme un processus aléatoire de loi inconnue, mais supposée identique quelle que soit la cellule. $EMG(t)$, représentant la bouffée d'activité résultant de la sommation de toutes ces activités élémentaires, peut donc être considéré comme un processus gaussien stationnaire par morceaux [duch93], [inba84]. Pour la suite il sera considéré comme modélisable par un processus AR à variance selon le segment analysé (activité de fond, bouffée).

Chaque signal $S_i(t)$ est un signal additif qui correspond à des activités différentes de l'EMG ou à des artefacts qui apparaissent à un instant donné. Ces signaux sont généralement courts par rapport au signal EMG(t). ils représentent les mouvements actifs fœtaux (MAF), les ondes d'Alvarez, les ondes LDBF (Longue Durée Basse Fréquence) et probablement d'autres types d'événements (figure 1.6).

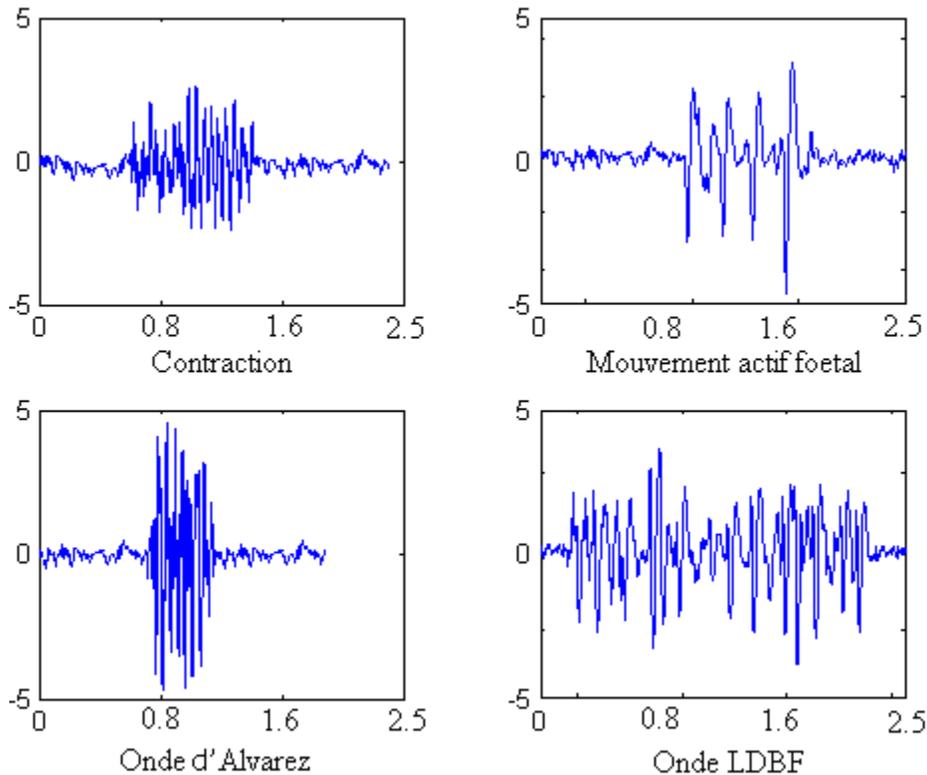


Figure 1.6. Les 4 types d'événements connus dans l'EMG utérin : une contraction, un mouvement actif fœtal, une onde d'Alvarez, et une onde LDBF. Axes des abscisses : unités en minutes. Axes des ordonnées : unités arbitraires.

Les instants d'apparition $S_i(t)$ sont imprédictibles. $S_i(t)$ peut donc dans certains cas se produire en même temps que EMG(t). La modélisation de tous ces événements n'est pas unique, mais leur distribution fréquentielle est un indicateur important permettant de distinguer entre ces événements. Leur apparition devait donc modifier l'erreur de prédiction (contenu fréquentiel et/ou variance) d'un modèle AR construit à partir de l'EMG seul.

Enfin, les bruits liés à l'environnement (principalement bruit d'électrodes et d'instrumentations) sont regroupés dans le terme $n(t)$. Dans la suite de ce travail, ce bruit sera supposé approximativement gaussien.

En première approximation, les contractions se présentent sous forme de bouffées d'activité très interférentielles alors que les artefacts de mouvement se présentent sous forme d'une succession des pics, traduisant chacun la réponse du système à la simulation mécanique induite par les mouvements actifs fœtaux. Les ondes d'Alvarez sont des contractions de durée beaucoup plus courte, de faible amplitude et de fréquence d'apparition élevée. Elles sont représentatives de contractions locales non propagées, survenant aléatoirement à différents endroits de l'utérus. Les ondes LDBF (Longue Durée Basse Fréquence) sont des événements de basse fréquence en comparaison au contenu fréquentiel des contractions au même terme.

1.5.3 Caractéristiques détaillées des événements contenus dans l'EMG utérin

Contractions

La fréquence d'apparition des contractions utérines varie beaucoup d'une femme à l'autre ainsi qu'en fonction du terme de grossesse [marq95]. Ce sont des contractions de longue durée (supérieure à une minute) et de fréquence d'apparition faible qui sont détectées sur toutes les femmes. La figure 1.7 représente deux contractions enregistrées à mi-grossesse et fin de grossesse ainsi que les densités spectrales de puissance associées. Cette figure montre la différence dans le contenu fréquentiel en fonction du terme de grossesse. Les contractions utérines présentent des caractéristiques spectrales de très basse fréquence (la quasi-totalité de l'énergie se situe à des fréquences inférieures à 0,5 Hz).

La contraction enregistrée à mi-grossesse présente une répartition d'énergie dans une bande de fréquence inférieure à 0,2 Hz alors que la contraction de fin de grossesse présente encore une énergie significative au delà de 0,3 Hz. 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. Elles apparaissent comme une succession de bouffées plus courtes, étalées dans le temps, alors que les contractions de grossesse plus tardives sont associées à une seule bouffée ininterrompue de durée inférieure à 2 min. Ceci peut être rapproché de l'évolution de la propagation de l'activité contractile utérine au cours du terme [cald60]. Les contractions peu propagées (de milieu de grossesse) induisent des bouffées plus longues dues à l'arrivée successive d'activités se propageant lentement à tout l'utérus. Au fur et à mesure de l'augmentation des capacités de propagation du muscle, ces différentes activités se regroupent sur un intervalle de temps plus court.

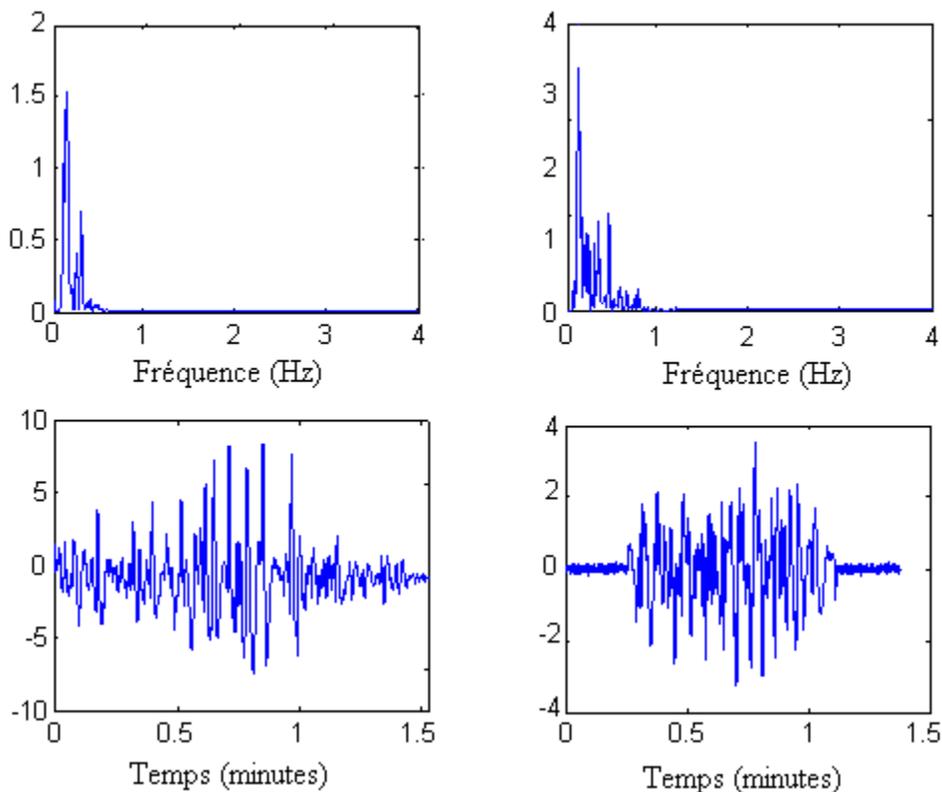


Figure 1.7. Gauche : contraction utérine à mi-grossesse et densité spectrale de puissance (DSP) associée. Droite : contraction utérine enregistrée fin de la grossesse et densité spectrale associée. Axes des ordonnées : unités arbitraires [marq95].

Caractéristiques fréquentielles des contractions

Pour aller plus loin dans le contenu spectral des contractions utérines, 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. De type biphasique, elle s'étale sur toute la durée de la contraction utérine. La seconde, appelée onde rapide, peut être enregistrée aussi bien en situation interne qu'externe, et pendant la grossesse et l'accouchement. Cette onde recouvre deux domaines spectraux 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(t). On tend à lier cette activité à des informations relatives à la propagation de l'activité électrique [deve95a]. L'activité FWH (pour Fast Wave High) est principalement observée dans des EMG correspond à des contractions efficaces d'accouchement. Son contenu spectral est directement relié à la fréquence des PA intra-bouffée. Il semble à l'origine du contrôle de l'intensité de la contraction. La figure 1.8 résume les connaissances sur le contenu fréquentiel de chacun de ces domaines dans le cas de la femme. Pour des signaux de contractions efficaces d'accouchement, Marque et al [marq89] ont observé un glissement du pic FWL vers les hautes fréquences, ainsi qu'une augmentation de l'énergie du pic FWH.

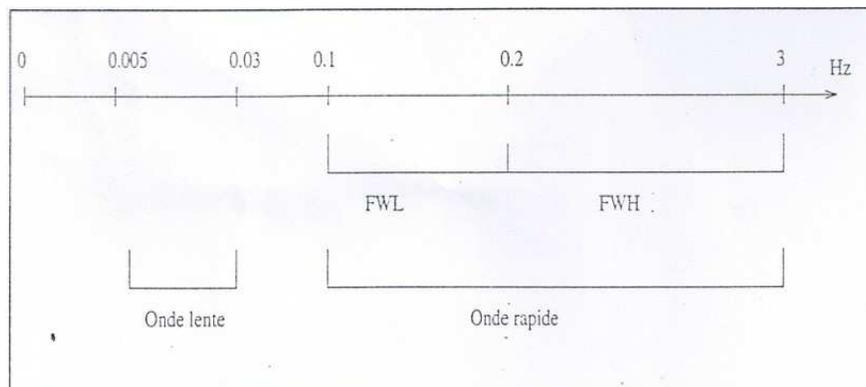


Figure 1.8. Caractéristiques du contenu fréquentiel d'EMG utérin chez la femme [deve95a].

Si on veut détecter des menaces d'accouchements prématurés et donc obtenir des informations liées à l'efficacité des contractions, il apparaît donc clairement qu'il faut s'intéresser au contenu spectral 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, médicaments divers) sont autant de paramètres riches de renseignements [deve95b], [duch95].

Mouvements actifs fœtaux (MAF)

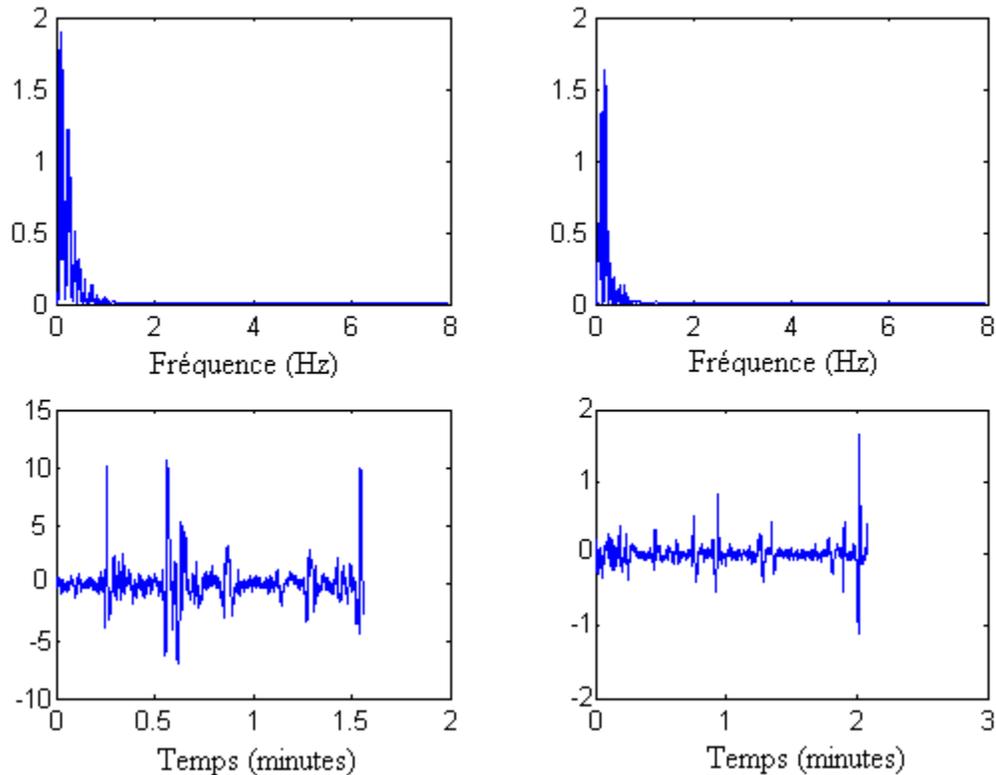


Figure 1.9. Deux exemples de mouvements actifs fœtaux. Gauche : MAF déclenchant une contraction avec la densité spectrale associée. Droite : MAF provoqué par une contraction avec la densité spectrale associée. Axes des ordonnées: unités arbitraires [marq95].

Les mouvements actifs fœtaux ne sont pas ressentis par les femmes à faibles termes de grossesse. Ces mouvements induisent en général des perturbations importantes sur le signal recueilli [marq95]. L'explication la plus évidente est qu'ils provoquent une déformation mécanique de l'abdomen dans la zone concentrée et donc des artefacts de type mouvement peau/ électrodes. Les artefacts liés à des mouvements actifs fœtaux sont rares à des termes antérieurs à 30 semaines d'aménorrhée (SA) (taille du fœtus inférieur à 30 cm pour 30 SA). La figure 1.9 représente deux types de mouvement actifs fœtaux avec leur densité spectrale associée. Le premier à gauche déclenche une contraction, le deuxième à droite est provoqué par des contractions. Notons que les mouvements actifs fœtaux peuvent être superposés avec les contractions et les ondes d'Alvarez.

Ondes d'Alvarez

Les ondes d'Alvarez sont des contractions de durée beaucoup plus courte (30 secondes à une minute) que les contractions classiques de grossesse. Leur fréquence d'apparition est élevée (une toutes les une à deux minutes).

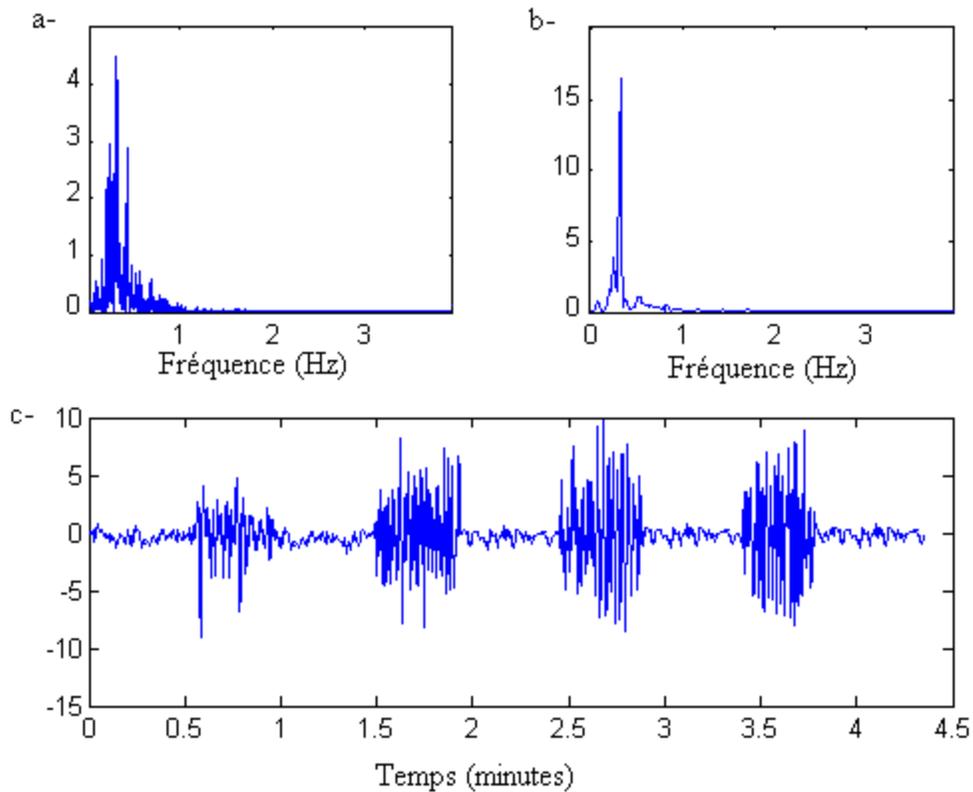


Figure 1.10. Ondes Alvarez identifiées dans un signal EMG utérin (c). (a) : DSP associée à l'ensemble de la fenêtre temporelle représentée en (c). (b) : DSP associée à la première bouffée de la fenêtre temporelle représentée en (c). Axes des ordonnées : unités arbitraire [marq95].

Elles sont représentatives de contractions locales non propagées survenant aléatoirement à différents endroits de l'utérus. Le contenu fréquentiel de ces bouffées (figure 1.10) est de plus haute fréquence (pic supérieur à 0,2Hz) que celui des contractions à un terme de grossesse équivalent. Généralement, elles sont associées soit à une activité physiologique normale, soit à une hyperactivité utérine susceptible d'évaluer ultérieurement vers un accouchement prématuré [newm87]. La zone d'investigation des électrodes utérines étant limitée par rapport à la taille totale du muscle, l'aspect local de ces ondes d'Alvarez peut expliquer le fait qu'elles n'aient été détectées que sur un petit nombre de femmes. Les électrodes ne peuvent en effet enregistrer une activité locale que si celle-ci provient d'une partie du muscle proche de la zone d'investigation.

Autres types

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'événements n'est pas explicitée.

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.

1.6 Traitement du signal EMG utérin

Le signal électromyographique a fait l'objet de nombreuses études depuis les années 50 [dill46], [plan84], [stee50], [sure56], [val79], [wolf79]. Jusqu'aux années 80, la majeure partie des auteurs ne se sont intéressés qu'à l'analyse du signal temporel et n'ont pu extraire d'information exploitable d'un point de vue clinique de l'analyse de ce signal.

Les bandes de fréquence permettant de segmenter correctement les bouffées d'activité électrique recueillies sur l'abdomen de la femme pendant l'accouchement ont été déterminées par val en 1979 [val79]. Marque en 1987 [marq87], a étudié la possibilité d'extraire de l'électromyogramme des informations pertinentes pour la surveillance de l'accouchement. L'EMG utérin était recueilli dans 2 situations de référence où l'efficacité des contractions est reconnue différente : fin de grossesse (contractions inefficaces) et accouchement (contractions efficaces). Les analyses temporelles et spectrales des signaux ont alors permis de décrire le signal au moyen d'un ensemble de trois paramètres représentatifs des deux classes prédéfinies : les énergies relatives dans 2 bandes de fréquence (H/T et B/T) et la durée de chaque bouffée d'activité. Une analyse discriminante, effectuée sur l'ensemble des individus des groupes de référence à partir de ces trois paramètres, a permis d'obtenir les vecteurs qui discriminent au mieux les deux situations observées. La projection des trois paramètres sur le premier vecteur discriminant a permis de concentrer l'information qu'ils contiennent pour définir un critère unique appelé 'critère d'efficacité', caractéristique d'une contraction. Le calcul en temps réel de ce critère a permis de suivre l'efficacité de chaque contraction tout au long du travail. Ce critère a ensuite été validé par l'ajout d'individus supplémentaires.

Cette équipe s'est ensuite attachée à valider l'utilisation de l'EMG utérin pour le suivi de la contractilité utérine pendant la grossesse. En effet, du fait du recueil abdominal indispensable à une surveillance obstétricale non invasive, l'origine utérine du signal recueilli avait été mise en doute. Le signal abdominal peut en effet contenir différentes sources d'artefacts (muscles abdominaux de la mère, mouvements respiratoires, ECG maternel, mouvements d'électrodes...), qui pourraient être responsables du manque de cohérence entre les résultats obtenus par l'analyse de ce signal [sure56], [nage83].

Pour cela, les caractéristiques de signaux, enregistrés pendant la grossesse chez la guenon par électrodes internes cousues sur l'utérus (signaux internes : EMG), ont été comparées aux caractéristiques des signaux enregistrés par électrodes abdominales positionnées en regard des électrodes internes (signaux externes EHG). Les caractéristiques temporelles et le contenu spectral des signaux ont ensuite été analysés, afin de mettre en évidence les relations existant entre les EMG et les EHG [mans93]. Les résultats prouvent que les deux types de signaux présentent des caractéristiques temporelles et spectrales similaires : mêmes instants d'apparitions des bouffées d'activité, contenus fréquentiels similaires.

Les deux types de signaux présentent en effet 2 bandes de fréquences caractéristiques, avec des énergies relatives dues aux différents modes de recueil. Ces deux bandes ont été identifiées comme FWL (Fast Wave Low : 0.1-1.2 Hz) pour la bande de plus basse fréquence, et FWH (Fast Wave High : 1.2-4.7 Hz) pour la bande de plus haute fréquence [deve93]. Une analyse plus fine fait apparaître, pour FWH, une suite de pics harmoniques que modélise bien un phénomène de modulation de fréquence à porteuse non sinusoïdale. Cette activité semble directement en relation avec l'activité cellulaire, représentée par des bouffées de potentiels d'action à fréquence variable [kuri61]. Une méthode de démodulation a été développée permettant de suivre l'évolution de la fréquence instantanée au cours de la contraction. Les résultats obtenus sont en accord avec les évolutions de l'activité cellulaire rapportées dans la littérature [mars84].

Les liens existants entre les caractéristiques du signal électromyographique et la physiologie de la contraction utérine ont ensuite été analysés afin d'essayer de comprendre les

phénomènes mis en jeu lors de la grossesse et de l'accouchement. Les principales modifications observées portent sur l'excitabilité cellulaire et sur des modifications de la propagation [kao77].

Des méthodes de type temps-fréquence ont alors été mise en place [deve95b]. Après analyse sur des signaux synthétiques, de l'outil temps-fréquence le plus adapté à l'étude de la modulante FWH, la modélisation AR glissante du signal a été sélectionnée. La forme de la modulante extraite de FWH a permis de mettre en évidence son évolution face aux différentes situations physiologiques explorées. Une forme de modulante typique (modulation 'en cloche') a été ainsi notée pour les contractions enregistrées à terme, aussi bien que pendant la grossesse après injection d'hormones excitatrices. Cette forme caractéristique de modulante a aussi été observée par d'autres auteurs au niveau cellulaire. Elle est associée à une activité mécanique intense du muscle utérin. Par ailleurs, les modulantes observées pour des contractions spontanées de grossesse ont une forme différente, voire aléatoire, associée à une activité mécanique beaucoup plus faible. Il semble donc exister une relation de cause à effet entre la forme de la modulante de l'activité FWH d'un EMG utérin et le caractère efficace de la contraction mécanique associée [deve93], [duch95], [deve95a].

La transformée d'ondelettes a déjà été utilisée par plusieurs auteurs pour le débruitage de l'EMG utérin [lema97], [lema98] ou pour suivre le contenu fréquentiel de ce signal [Khal97]. En 1997, Buhimschi a confirmé que les paramètres fréquentiels peuvent être utilisés pour caractériser l'activité électrique de manière non invasive [bui97]. En 1998 Garfield et al. ont confirmé que l'activité myométriale peut être détectée sur la surface abdominale [garf98]. La même année, Buhimschi et al. ont montré eux aussi que l'analyse de la puissance spectrale peut être utilisée pour quantifier les signaux EMG utérins [bui98].

Khalil en 1999, a présenté une méthode originale permettant la détection et la classification de différents événements présents dans un signal non stationnaire. La détection et l'isolement des différents événements contenus dans le signal enregistré est une étape essentielle pour l'utilisation de l'EMG utérin dans le cadre de la Prévention des Menaces d'accouchement Prématuro (MAP).

Cette méthode utilise deux techniques différentes de détection, l'une fondée sur une approche monodimensionnelle, l'autre sur une approche multidimensionnelle. Ce second 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 [Khal99].

1.7 Détection des MAP par classification des EMG utérins

En 2001 Garfield et al. ont montré que toutes les méthodes utilisées couramment pour le traitement ou la prévention de l'accouchement prématuré comme le tocodynamomètre, la fibronectine foetale, les ultrasons, possèdent une faible capacité pour la prédiction de l'accouchement, en comparaison avec l'analyse de l'EMG utérin, pour lequel ils ont démontré un pourcentage de prédiction très élevées [garf01].

Maner et al en 2003 ont travaillé sur un total de 99 patientes. Ces patientes sont divisées en deux groupes : groupe d'accouchement à terme (plus que 37 Semaines d'Aménorrhées, SA, n=57), groupe d'accouchement prématuré (moins que 37 SA, n=42). L'activité électrique utérine a été enregistrée pendant 30 minutes. L'EMG utérin a été traité pour estimer sa densité spectrale de puissance (DSP). Le délai terme à la mesure - terme à l'accouchement a été comparé à la fréquence maximale du spectre de la DSP. Les courbes ROC ont été

calculés pendant 48, 24, 12 et 8 heures du groupe à terme, et 6, 4, 2 et 1 jours pour le groupe prématuré [mane03].

Cette étude a montré que la fréquence maximale du spectre de la DSP augmente quand le délai terme à la mesure - terme à l'accouchement diminue. L'analyse des courbes ROC a donné des valeurs prédictives positives et négatives élevées pour l'accouchement à terme et l'accouchement prématuré. À la limite, la fréquence maximale du spectre de la DSP des femmes du groupe d'accouchement à terme est la plus élevée pour le sous-groupe des signaux enregistrés 24h avant l'accouchement [mane03]. Pour le groupe des accouchements prématurés, la fréquence maximale de la DSP est la plus élevée pour le sous-groupe des signaux enregistrés mais de 4 jours avant l'accouchement [mane03]. Maner et al. ont donc conclu que l'accouchement à terme peut être détecté 24 heures à l'avance, et que l'accouchement prématuré peut être détecté 4 jours avant la date d'accouchement.

En 2007 Maner et al. ont refait la même analyse en augmentant la cohorte de patientes à 185, 134 femmes ayant accouché à terme et 51 femmes prématurément [mane07]. Ces patientes sont divisées en quatre sous-groupes : en travail à terme (TL, n=75), en travail pré-terme (PTL, n=13), à terme non en travail (TN, n=59), avant terme non en travail (PTN, n=38). Les paramètres de classifications utilisées sont, la fréquence maximale de chaque DSP calculée à partir de la bouffée d'EMG utérin associée à la contraction utérine, la durée de la contraction, le nombre de contractions par unité de temps, ainsi que le temps total de contraction. Ils ont utilisé les réseaux de neurones artificiels comme outil de classification, Ils ont obtenus 59/75 (79%) de patientes bien classifiées en TL, 12/13 (92%) en PTL, 51/59 (86%) en TN, 27/38 (71%) en PTN [mane07].

Dans ce contexte de détection et de classification des accouchements prématurés Verdenik et al. [verd01] ont travaillé sur les signaux EMG utérins de 47 patientes enregistrées entre 25 et 35 SA. 30 femmes (63,8%) ont accouché après 37 semaines de grossesse, et 17 femmes (36,2%) ont accouché prématurément [verd01]. Ils ont utilisé la RMS (Root Mean Square) et la fréquence médiane de la DSP comme paramètres fréquentiels pour caractériser les EMG utérins. Ils ont aussi utilisé des paramètres d'ordre général caractérisant les patientes comme l'âge de la patiente, son poids à l'admission, fumeuse ou non, l'histoire gynécologique des anomalies utérines, la parité, les avortements précédents (spontané ou induit), les accouchements prématurés précédents, le terme en semaines d'aménorrhées à l'admission, la présence de contractions à son arrivée à l'hôpital, le traitement avec des agents spasmolytique et/ou tocolytique. Parmi les 14 variables indépendants analysés dans ce travail, seuls la RMS et le poids possèdent une corrélation positive avec l'accouchement prématuré. Ces conclusions ont aussi été trouvées par Carmichael et al. [carm97]

La corrélation entre les paramètres de l'EMG et l'âge gestationnel montre qu'avec l'augmentation du terme, le contenu fréquentiel de l'activité électrique devient plus basse fréquence, ce qu'exprime une fréquence médiane plus basse de l'activité électrique utérine. En même temps la valeur RMS n'est pas significativement corrélée avec l'âge gestationnel.

Ces résultats sont opposés à ceux de Buhimschi et al. [buhi96], [buhi98], qui ont constaté que le contenu spectral de l'EMG utérin augmente quand on s'approche de l'accouchement. Cette différence est liée au fait que Vedrenik [vedr01] a travaillé non pas sur chaque bouffée d'EMG utérin associée aux contractions (souvent inexistantes aux termes précoces de grossesse), mais sur des intervalles de 30 minutes qui incluent la ligne de base du signal (souvent bruité dans les hautes fréquences) et les éventuelles bouffées d'activité utérine (de fréquence plus basse que le bruit de la ligne de base). Les bruits électromagnétiques environnants ainsi que le 50 Hz (ou 60 Hz) contaminent, dans la plupart des cas, le signal EMG utérin. La deuxième explication possible de cette différence est que Buhimschi et Garfield ont utilisé un filtre passe-bande 0.3-50 Hz [buhi96], tandis que Vedrenik a choisi un filtre passe-bande 0.1-4 Hz.

Leman en 1999 a défini une liste de paramètres utilisables pour séparer la population des accouchements à terme (AT) de celles ayant accouchés prématurément (AP) [lema99]. Ces paramètres (91), dont les plupart sont extraits des arêtes du scalogramme de l'EMG utérin

et de l'étude de ses zones de stationnarité, justifiés par des études précédentes [deve93], étant trop nombreux par rapport au nombre de signaux pour être utilisés comme entrée d'un système de classification, elle a été amenée à éliminer les paramètres corrélés par analyse de la matrice de corrélation et classification ascendante hiérarchique. Elle n'a finalement retenu que 57 paramètres. Puis elle a classifié les signaux EMG utérins selon la position du placenta (postérieur, antérieur) en utilisant l'analyse factorielle discriminante. Cette étude a montré qu'il est possible de séparer linéairement les EMG utérins Accouchement Prématuré (AP)/Accouchement à Terme (AT), enregistrés avant 28 SA, pour les femmes possédant un placenta postérieur, avec un taux de mauvais classement acceptable (17%). Pour la population qui possède un placenta antérieur, la séparation est possible par discrimination linéaire avant 28 SA mais avec un taux de mauvais classement moins bon (32-35%).

Dans cette même équipe, Voisine en 2001, après avoir développé un prototype d'appareillage pour l'enregistrement à domicile de l'EMG utérin pendant la grossesse [vois00], a étudié la possibilité de classer les EMG utérins entre AT et AP [vois01]. Elle a utilisé les mêmes paramètres que Leman [vois01], mais avec des réseaux de neurones de type Perceptron Multi Couches (PMC) comme outil de classification. Deux types de PMC utilisant deux algorithmes d'apprentissage différents, ont été analysés et comparés : le premier est basé sur l'algorithme de descente du gradient avec inertie (Back Propagation with Momentum, BPM), le deuxième est basée sur l'algorithme de Levenberg-Marquart (LM) avec une régulation bayésienne. Les taux de mauvais classement obtenus sur l'ensemble des données testées sont présentés au tableau 1.1 [vois01].

Placenta	<28 SA		28-32 SA		33-37 SA	
	Post.	Ant.	Post.	Ant.	Post.	Ant.
BPM	18%	-	76%	26%	19%	19%
LM	13%	-	19,5%	18%	12%	11%

Tableau 1.1. Taux de mauvais classement obtenu pour les méthodes testées pour la détection des MAP pour les deux algorithmes, BPM (Back Propagation with Momentum) et LM (Levenberg-Marquart), selon le terme et la position du Placenta (Post : placenta postérieur, Ant : placenta antérieur). SA = semaine d'aménorrhée.

Elle a de plus comparé la Sensibilité (Se), la Spécificité (Sp) et la Valeur Prédicative Négative (NPV) des deux méthodes [vois02] (Tableau 1.2), en différenciant, comme Leman, les cas de placenta antérieurs et postérieurs.

Placenta	BPM		LM	
	Post.	Ant.	Post.	Ant.
Se	0,85	0,9	0,82	0,79
Sp	0,93	0,91	0,88	0,87
NPV	0,88	0,92	0,87	0,85

Tableau 1.2. Performances des méthodes utilisées pour la détection des MAP pour les deux algorithmes, BPM (Back Propagation with Momentum) et LM (Levenberg-Marquart), selon la position du Placenta (Post : placenta postérieur, Ant : placenta antérieur). Se : Sensibilité, Sp : Spécificité, NPV : Valeur Prédicative Négative

Voisine conclut que l'algorithme LM fournit de meilleurs résultats que le BPM, en termes de taux de mauvais classement, ainsi qu'en termes de contractions non classées (3-45% pour BPM, <2,5% pour LM). Cependant, si on compare les Sensibilité(Se), Spécificité (Sp) et Valeur Prédicative Négative (NPV), les résultats sont légèrement meilleurs pour le BPM quel que soit le terme. Elle n'a pas, comme Leman, mis en évidence de meilleurs résultats de classification pour les placentas postérieurs. Elle signale de plus, que le principal problème demeure le faible nombre de contractions disponibles pour l'apprentissage ou le test, ce qui ne favorise pas une bonne généralisation de l'algorithme [vois02].

1.8 But de cette thèse

Après les premières tentatives de détection des MAP à l'aide d'outils de classification supervisée réalisés par Leman [lema99] et Voisine [vois01], notre travail a été orienté vers la classification des contractions dans les signaux EMG utérins afin de séparer les deux types d'accouchement, en utilisant d'autres méthodes supervisées et non supervisées.

Après l'isolation des contractions utérines à partir du signal EMG utérin, chaque contraction est modélisée et des paramètres sont extraits avant de faire la classification. Cette modélisation est faite par ondelette ou par analyse de la densité spectrale de puissance de chaque contraction.

Ensuite, les algorithmes de classifications supervisées comme les réseaux des neurones à fonction radiaux (RBF, Radial basis function), et les réseaux d'ondelettes (Wavnet, Wavelet network) sont utilisés pour identifier les contractions.

De plus pour identifier les contractions des méthodes de classification non supervisées comme les réseaux de neurones compétitifs et la méthode de classification non supervisée basée sur le test de Fisher combiné avec la méthode de k-moyenne (USCM, Unsupervised Statistical Classification Method) sont appliqués pour faire une classification ou clustering.

1.9 Conclusion

Ce chapitre a été consacré à une revue des données physiologiques nécessaires d'une part à la bonne compréhension des événements en jeu, d'autres parts à une définition claire du problème. Tout d'abord nous avons expliqué le MAP (Menace d'accouchement prématuré), ses causes, et les méthodes utilisées en gynécologie-obstétrique pour détecter et prévenir les accouchements prématurés (Mesure de marqueurs biochimiques, Signes cliniques, La détection du rythme circadien de l'activité utérine...).

Le signal EMG semble être un vecteur potentiel d'indication du risque d'accouchement prématuré. Nous avons décrit ce signal EMG dans ce chapitre. Il est apparu clairement que l'EMG utérin est un signal non stationnaire dont les composantes semblent être riches en renseignements quant à la nature efficace de la contraction correspondante. Les caractéristiques temporelles et fréquentielles de tous les événements ont été expliquées. A la fin de ce chapitre nous avons introduit le but de cette thèse qui peut être résumé comme une classification entre contractions à termes ou prématurés. En utilisant des méthodes de classification supervisées et non supervisées qui n'avaient pas été testées jusqu'alors pour la détection des MAP.

Chapitre 2

Modélisation des EMG utérins et extraction des paramètres

2.1 Introduction

L'extraction des paramètres dans les signaux est une étape primordiale avant la classification. Toute méthode de classification nécessite une sélection des paramètres pertinents, discriminants et les plus adaptés au signal.

Les paramètres liés à un signal aléatoire sont généralement statistiques ou fréquentiels. Les paramètres statistiques sont calculés à partir de la densité de probabilité. Les paramètres fréquentiels dépendent de la densité spectrale de puissance. D'autres paramètres importants peuvent être extraits de décomposition en ondelettes.

Dans ce chapitre, nous présentons tout d'abord la décomposition en ondelettes discrètes et l'analyse multirésolution. A la fin de cette section nous décrivons les paramètres extraits à partir des coefficients d'ondelettes.

La modélisation Autorégressive est un outil important en traitement du signal. Les paramètres AR donnent des informations sur le contenu fréquentiel du signal. Ils peuvent aussi être utilisés pour reconstruire des signaux. La section 2 de ce chapitre présente donc la modélisation AR, le choix de l'ordre de modèle et la méthode d'estimation de la densité spectrale à partir de ce modèle.

Nous présenterons ensuite d'autres méthodes d'estimation de la DSP. Les paramètres extraits de cette représentation seront présentés à la fin de ce chapitre.

2.2 La transformée en ondelette

Le terme *ondelette* désigne une fonction qui oscille pendant un " temps donné" (si la variable est le temps) ou sur un intervalle de longueur finie (si la variable est de type spatial). Au delà, la fonction décroît très rapidement vers zéro.

Historiquement, les premières ondelettes (introduites par Haar dans les années 1930) constituaient une base de fonctions orthogonales. Les ondelettes de Haar présentent la particularité de ne pas être dérivables.

Plus récemment, de nouvelles fonctions ondelettes ont été introduites, qui constituent également une base de fonctions orthogonales, et qui, de plus, sont dérivables. Elles ont été notamment mises en œuvre dans le cadre de l'analyse multirésolution de signaux et par suite, dans leur classification. Ces ondelettes ne peuvent s'exprimer sous une forme analytique simple. Pour cette raison, elles sont peu adaptées pour servir notre objectif. Nous n'utiliserons donc pas ce type d'ondelettes dans ce mémoire.

Les structures obliques (*frames* en anglais) ont été introduites par J. Morlet dans le but de trouver des bases de fonctions (non nécessairement orthogonales) pour représenter des

signaux. Ces structures obliques ont été ensuite l'objet des travaux de Daubechies, qui a développé un support théorique aux résultats de Morlet. Les structures obliques ont des expressions analytiques simples, et toute fonction de carré sommable peut être approchée, avec la précision voulue, par une somme finie d'ondelettes issues d'une structure oblique. Pour toutes ces raisons, nous nous sommes intéressés uniquement, dans nos applications de classification, à des structures obliques d'ondelettes.

De manière analogue à la théorie des séries de Fourier, les ondelettes sont principalement utilisées pour la décomposition de fonctions. La décomposition d'une fonction en ondelettes consiste à l'écrire comme une somme pondérée de fonctions obtenues à partir d'opérations simples effectuées sur une fonction principale appelée *ondelette-mère*. Ces opérations consistent en des translations et dilatations de la variable. Selon que ces translations et dilatations sont choisies de manière continue ou discrète, on parlera d'une transformée en ondelettes continue ou discrète.

2.2.1 Transformée en ondelettes continues

La transformée en ondelettes permet d'analyser efficacement des signaux ou se combinent des phénomènes d'échelles très différentes. La transformée en ondelettes continue d'un signal $x(t)$ prend la forme [flan93], [krim95] :

$$T_x^\psi(a, b) = \int_{-\infty}^{+\infty} x(t) \psi_{ab}(t) dt \quad 2.1$$

où

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

est l'ondelette analysante. Elle est localisée en temps et en fréquence. Cette équation correspond à traduire une ondelette de référence après l'avoir dilatée ($a > 1$) ou comprimée ($a < 1$). Donc en variant les paramètres a (dilatation ou compression de l'ondelette) et b (translation temporelle), on balaye l'ensemble de l'axe des temps et des fréquences. La signification des coefficients d'ondelette $T_x^\psi(a, b)$ peut être décrite à partir des trois remarques suivantes :

- l'ondelette doit être localisée en temps, cela signifie qu'en dehors d'un intervalle $[0, L]$ la fonction ψ prend des valeurs négligeables : ψ est concentrée sur l'intervalle $[0, L]$.

- ψ doit aussi être localisée en fréquence, son spectre est concentré sur une bande de fréquences $B = [f_1 ; f_2]$, $0 < f_1 < f_2$.

- la fonction $\psi_{ab}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$ est localisée sur l'intervalle de temps $[b ; b+L.a]$. Sa

transformée de Fourier est localisée sur l'intervalle $\left[\frac{f_1}{a} ; \frac{f_2}{a}\right]$. La résolution temporelle

correspondant au coefficient $T_x^\psi(a, b)$ est L.a, la résolution fréquentielle est

$$\frac{f_1 - f_2}{a} = \frac{B}{a}.$$

Donc la transformée en ondelettes peut s'interpréter comme une analyse par banc de filtres,

à surtension constante, $Q = \frac{f_0}{B}$. f_0 est la fréquence centrale de l'ondelette prise pour référence à l'échelle 'naturelle' $a=1$.

Il est possible de reconstruire le signal $x(t)$ à partir de sa décomposition en coefficients d'ondelette :

$$x(t) = A \iint T_a^\psi(a, b) \psi\left(\frac{t-b}{a}\right) da db \quad 2.3$$

A est une constante de normalisation.

Les conditions d'admissibilité de la transformée en ondelettes sont:

$$\int_{-\infty}^{+\infty} |\Psi(f)|^2 \frac{df}{|f|} = 1 \quad 2.4$$

L'interprétation physique en est simple. En effet, afin que l'intégrale ci-dessus soit finie, il faut s'assurer de sa convergence tant à l'infini qu'à l'origine.

Une deuxième condition est plus sévère, puisqu'elle impose que le spectre de $\psi(t)$ s'annule convenablement à l'origine, de façon à prévenir une divergence possible due à la mesure d'intégration en $|f|^{-1}$. Cette deuxième condition impose en particulier que $\psi(t)$ soit a moyenne nulle :

$$\int_{-\infty}^{+\infty} \psi(t) dt = \Psi(0) = 0 \quad 2.5$$

et qu'ainsi elle présente au moins quelques oscillations, d'où le nom d'ondelette.

2.2.2 Transformée en ondelettes discrètes

Les paramètres d'échelle a et de translation b intervenant dans la définition des coefficients d'ondelettes sont définis comme étant continus. Dans le cadre de ce travail, nous sommes concernés par une décomposition multiéchelles (transformée en ondelettes discrètes), donc par des valeurs discrètes de a et b .

Nous donnons ci-dessous une version discrétisée de la transformation en ondelettes en posant :

$$- a = a_0^m \text{ avec } a > 1$$

$$- b = n.b_0.a_0^m \text{ avec } b_0 > 0, \text{ avec } m \text{ et } n \text{ entiers.}$$

Les fonctions analysantes ψ_{ab} seront notées dans ce cas $\psi_{mn}(t) = \psi(a_0^m t - n b_0)$; les coefficients d'ondelettes seront notés $T_x^\psi(m, n)$ au lieu de $T_x^\psi(a, b)$. Le coefficient $|T_x^\psi(m, n)|^2$

représente donc l'énergie du signal dans l'intervalle de temps $[k.b_0.a_0^m; (k+1)b_0.a_0^m]$ et la bande de fréquences $\left[\frac{1}{a_0^{m+1}}, \frac{1}{a_0^m} \right]$. La formule de décomposition d'un signal $x(t)$ en coefficients discrets d'ondelette ne change pas par rapport à la définition continue. En effet nous prenons simplement des valeurs particulières de a et b :

$$T_x^\psi(n, m) = \int_{-\infty}^{+\infty} x(t) \psi_{nm}(t) dt \quad m, n \in \mathbb{Z} \quad 2.6$$

La formule de reconstruction permettant de retrouver $x(t)$ à partir de l'ensemble discret de ses coefficients d'ondelette devient :

$$x(t) = \sum_{m,n} T_x^\psi(n, m) \psi_{nm}(t) \quad m, n \in \mathbb{Z} \quad 2.7$$

Notons que si $a_0=2$ la décomposition s'appelle décomposition dyadique en ondelettes.

2.2.3 Analyse multirésolution

L'obtention de bases discrètes d'ondelettes et plus particulièrement de bases orthonormées, repose sur la théorie de l'analyse multirésolution. Quantitativement, cette théorie prévoit que l'on peut analyser un signal en le décomposant en un signal d'approximation et un signal de détail. Les approximations sont en quelque sorte des versions lissées passe-bas du signal, dans lesquelles les ruptures sont progressivement gommées ou atténuées. Les détails ressemblent au squelette du signal, ne comportant que les informations de ruptures ou discontinuités [abry97].

Définition 1 une analyse multi-résolution est une suite $\{V_m\}, m \in \mathbb{Z}$ de sous espaces de $L^2(\mathbb{R})$ vérifiant les propriétés suivantes :

- $V_{m+1} \subset V_m$: les V_m sont emboîtés les uns dans les autres.
- $\bigcap_m V_m = L^2(\mathbb{R})$: pour toute fonction de $L^2(\mathbb{R})$ on peut trouver au moins un des ensembles V_m qui la contient.
- $g(\cdot) \in V_m \Leftrightarrow g(2\cdot) \in V_{m+1}$ si la fonction g appartient à V_m , la même fonction dilatée d'un rapport 2 appartient à V_{m+1} .
- $g(\cdot) \in V_0 \Leftrightarrow g(\cdot - n) \in V_0$ si la fonction g appartient à V_0 alors la même fonction translatée de n est encore dans V_0 .
- de plus il doit exister une fonction $\phi(t)$ telle que l'ensemble des fonctions $\{ \phi_{nm}(t) = 2^{m/2} \phi(2^m t - n) \}$ constitue une base orthonormée de V_0 . Cette fonction ϕ est appelée fonction échelle.

Les espaces V_m sont appelés espaces d'approximation. De plus si $V_k \subset V_m$, la projection dans V_k d'un signal $x(t)$ de $L^2(\mathbb{R})$ constitue une approximation plus grossière que la projection V_m . Les projection successives de $x(t)$ dans V_m constituent autant d'approximations du signal à des niveaux de résolution différents et contiennent de moins en moins d'informations relatives à x .

Définition 2 l'analyse multirésolution est r -régulière ($r \in \mathbb{N}$) s'il est possible de choisir la fonction ϕ_0 de sorte que l'on ait :

$$\left| \frac{d^k \phi_0(t)}{dt^k} \right| \leq C_m (1+|t|)^{-m} \quad 2.8$$

Pour $0 \leq k \leq r$, et tout entier $m \geq 1$ [abry97]. Cette requête de régularité caractérise à la fois des propriétés de bonne localisation temporelle et de 'douceur' de la forme de l'ondelette.

L'idée de l'analyse multirésolution est d'effectuer des projections successives du signal à étudier dans les V_m pour en obtenir des approximations :

$$approx_m(t) = (Proj_{V_m} x)(t) \quad 2.9$$

Comme V_m est inclus dans V_{m-1} , $approx_m(t)$ constitue une approximation de $x(t)$ plus grossière que $approx_{m-1}(t)$. L'information perdue entre deux approximations successives est appelée détail :

$$détail_m(t) = approx_{m-1}(t) - approx_m(t) \quad 2.10$$

La suite d'une analyse multirésolution consiste à examiner la suite des signaux d'approximations : $approx_m(t)$.

Une des conséquences de l'analyse multirésolution est que les signaux de détails peuvent être obtenus par projection du signal $x(t)$ dans un ensemble d'espaces $\{W_m, m \in \mathbb{Z}\}$, appelés espaces d'ondelettes :

$$détail_m(t) = (Proj_{W_m} x)(t) \quad 2.11$$

W_m est un complémentaire de V_m dans V_{m-1} ,

$$V_m + W_m = V_{m-1} \quad 2.12$$

Cette analyse multirésolution montre qu'il est possible de construire à partir de la fonction d'échelle ϕ_0 , une fonction ψ_0 appelée ondelette mère, telle que les fonctions $\left\{ \psi_{nm}(t) = 2^{m/2} \psi(2^m t - n) \right\}$ constituent une base orthonormée de W_m .

On peut déduire finalement que l'analyse multirésolution consiste à décomposer le signal $x(n)$ en utilisant l'ondelette ψ et sa fonction d'échelle ϕ . A l'aide de ces deux fonctions, on peut décomposer le signal en détails et approximations, respectivement. Les formules de calcul des coefficients des approximations et des détails peuvent s'écrire :

$$a_x(m, n) = \int_{-\infty}^{+\infty} x(t) \phi_{mn}(t) dt \quad 2.13$$

$$d_x(m, n) = \int_{-\infty}^{+\infty} x(t) \psi_{mn}(t) dt \quad 2.14$$

La reconstruction des approximations et des détails sur chaque échelle m est déduite des formules suivantes :

$$détail_m(t) = \sum_n d_x(m, n) \psi_{mn}(t) \quad 2.15$$

$$approx_m(t) = \sum_n a_x(m, n) \phi_{mn}(t) \quad 2.16$$

Fondamentalement, l'analyse multirésolution réécrit donc un signal sous forme d'une somme de détails et d'une ultime approximation à un niveau d'échelle J :

$$\begin{aligned} x(t) &= approx_J(t) + \sum_{m=1}^J detail_m(t) \\ &= \sum_n a_x(J, n) \phi_{Jn}(t) + \sum_{m=1}^J \sum_n d_x(m, n) \psi_{mn}(t) \end{aligned} \quad 2.17$$

2.2.4 Filtres de l'ondelette et de la fonction d'échelle

La famille des fonctions $\{\phi_{nm}(t) = 2^{m/2} \phi(2^m t - n)\}$ constitue une base orthonormée de V_0 . Alors si la fonction d'échelle $\phi(t)$ est dans V_0 , elle est également dans V_1 , et il existe nécessairement une famille de coefficients $h(n)$ tels que l'on puisse écrire :

$$\phi(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} h(n) \phi(2t - n) \quad 2.18$$

avec

$$h(n) = \sqrt{2} \int_{-\infty}^{+\infty} \phi(t) \phi(2t - n) dt; \sum_{n=-\infty}^{+\infty} h^2(n) = 1 \quad 2.19$$

Parce que les translatées entières de la fonction d'échelle $\phi(t)$ forment une base de V_0 , le filtre $h(n)$ possède des propriétés très particulières :

$$|H(f)|^2 + |H(f + 1/2)|^2 = 1 \quad 2.20$$

$$\sum_{n=-\infty}^{+\infty} |\Phi(f + n)|^2 = 1 \quad 2.21$$

avec $H(f)$ et $\Phi(f)$ les transformées de Fourier de $h(n)$ et $\phi(t)$, respectivement.

Les fonctions d'ondelettes $\{\psi_{nm}(t) = 2^{m/2} \psi(2^m t - n)\}$ constituent une base orthonormée de W_m . On peut également caractériser l'ondelette $\psi(t)$ par un filtre discret de coefficients $g(n)$. Alors comme $\psi(t)$ est inclus dans W_0 , lui-même inclus dans V_1 on peut écrire :

$$\psi(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} g(n) \phi(2t - n) \quad 2.22$$

avec

$$g(n) = \sqrt{2} \int_{-\infty}^{+\infty} \psi(t) \phi(2t - n) dt; \sum_{n=-\infty}^{+\infty} g^2(n) = 1 \quad 2.23$$

Dans la théorie des bancs de filtres, les filtres discrets $h(n)$ et $g(n)$ forment une paire de filtres miroirs en quadrature. Dans ce cas la relation entre g et h peut s'écrire [flan93] :

$$g(n) = (-1)^n h(1-n) \quad 2.24$$

Pour résumer, en partant de la fonction d'échelle $\phi(t)$ telle que la famille $\{\phi(t-n); n \in \mathbb{Z}\}$ soit une base de V_0 , et des coefficients $h(n)$ du filtre discret qui lui est associé, l'ondelette :

$$\psi(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} (-1)^n h(1-n) \phi(2t-n) \quad 2.25$$

est telle que la collection $\{\psi_{nm}(t) = 2^{m/2} \psi(2^m t - n)\}$ est une base de $L^2(\mathbb{R})$.

2.2.5 Propriétés et caractéristiques des ondelettes orthogonales

Les ondelettes ont quelques caractéristiques telles que la localisation, la compacité du support,..., etc. Nous 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:

$$\forall n = 0, \dots, r, \left| \frac{d^n}{dt^n} \psi_0(t) \right| \leq C_n \quad 2.26$$

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:

$$\forall k \in \{0, \dots, N-1\}, \int t^k \psi(t) dt = 0 \quad 2.27$$

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

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 [abry97]. 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.

2.2.6 Extraction des paramètres

Après avoir fait la décomposition en ondelettes, le but consiste maintenant à extraire les paramètres. Pour cela nous avons choisi les variances sur les échelles (variances de détails calculées sur les échelles, figure 2.1). Ce choix repose sur des travaux faits lors de la thèse de Khalil M. en 1999 [khal99]. Le choix des échelles et le nombre de paramètres sont discutés ultérieurement.

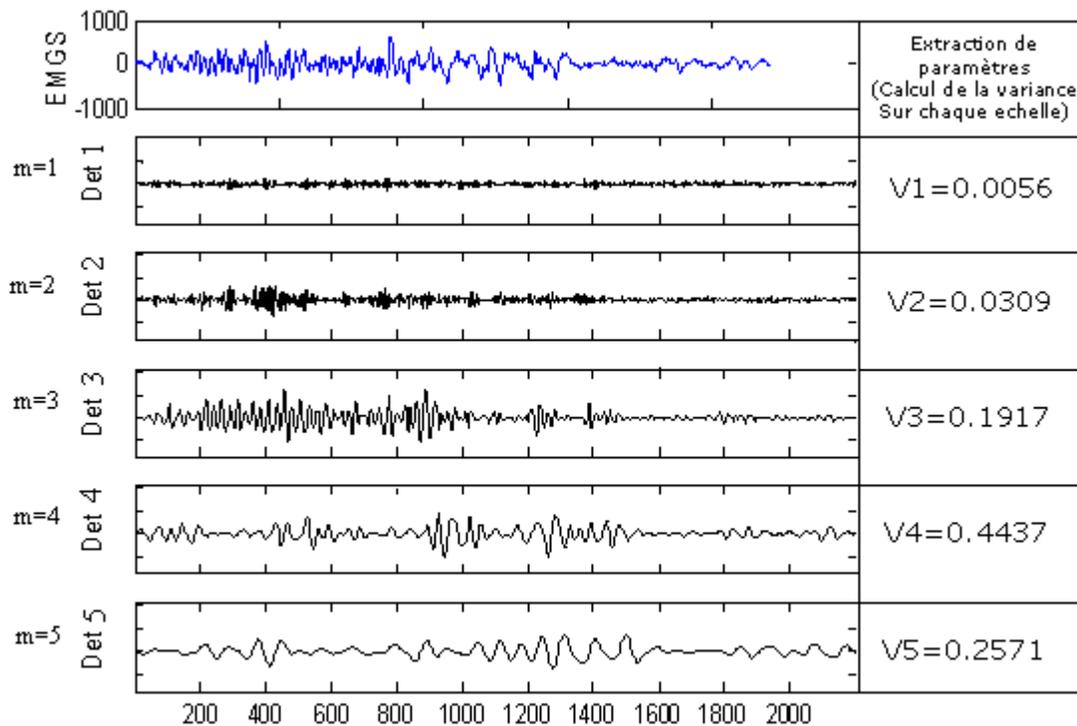


Figure 2.1. Décomposition par Ondelette 'Sym5' sur 5 niveaux d'échelle. La colonne de droite montre les variances de détails calculées sur les échelles.

2.3 Modèle AR

2.3.1 Les méthodes d'estimation des paramètres AR

Dans ce paragraphe nous présentons rapidement les méthodes les plus connues de modélisation AR. Cette méthode sera utilisée pour l'apprentissage de nos algorithmes de classification. En effet, du fait du faible nombre de signaux dans notre base de données, la modélisation AR a été utilisée pour créer des signaux synthétiques semblables aux signaux EMG utérins. Cette méthode a permis de créer une base de données synthétique de taille importante puis de tester les méthodes de classification sur cette base. Mais avant de créer les signaux synthétiques, il faut estimer les paramètres AR. Le choix de la méthode d'estimation influe sur l'algorithme de classification lui-même, ce qui nous a conduit à faire une comparaison des différentes méthodes d'estimation des paramètres AR afin de choisir la plus pertinente vis à vis de la classification des signaux EMG.

Notons aussi que les paramètres AR, les a_i , vont être utilisés pour estimer la densité spectrale du signal EMG.

Méthode de Yule-Walker

La méthode de Yule-Walker (dite aussi méthode d'autocorrélation) repose sur le remplacement de la minimisation statistique de l'erreur de prédiction par une minimisation de cette erreur sur l'ensemble des données x_0, \dots, x_{N-1} . Il s'agit donc de minimiser la somme :

$$\sum_{n=0}^{N+p-1} \varepsilon_p^2(n) = \sum_{n=0}^{N+p-1} (x(n) + \sum_{k=1}^p a_k \cdot x(n-k))^2 \quad 2.28$$

où ε est l'erreur de prédiction, les a_k sont les coefficients du modèle, p est l'ordre du modèle.

Ceci conduit à l'équation matricielle suivante [kay88] :

$$\begin{pmatrix} \hat{R}_{xx}(0) & \hat{R}_{xx}(1) & \dots & \hat{R}_{xx}(p) \\ \hat{R}_{xx}(1) & \hat{R}_{xx}(0) & \dots & \hat{R}_{xx}(p-1) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \hat{R}_{xx}(p) & \hat{R}_{xx}(p-1) & \dots & \hat{R}_{xx}(0) \end{pmatrix} \begin{pmatrix} 1 \\ a_1 \\ \cdot \\ \cdot \\ a_p \end{pmatrix} = \begin{pmatrix} \hat{\sigma}_\varepsilon^2 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix} \quad 2.29$$

avec

$$\hat{R}_{xx}(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} x(n) \cdot x(n+k); \quad k = 0, 1, \dots, p \quad 2.30$$

La méthode consiste donc à remplacer la matrice de corrélation dans l'équation de Yule-Walker normale par son estimateur biaisé [kay88]. La matrice reste une matrice de Toeplitz définie positive et l'algorithme de Levinson peut ainsi s'appliquer pour la résolution du système d'équations linéaires. Une discussion sur l'efficacité de cette méthode sera proposée par la suite.

Méthode de Burg

L'idée de Burg [kay88] fut de calculer directement à partir des données une estimée des coefficients de réflexion (équivalent aux coefficients d'AR) ce, sans passer par le calcul des covariances.

On peut calculer les coefficients de réflexion \hat{K}_p d'une façon récursive sur l'ordre, puis calculer les paramètres AR par la relation de Levinson :

$$\left\{ \begin{array}{l} \hat{K}_p = -(\sum_{i=0}^{p-1} \hat{R}(p-i) \cdot \hat{a}_{p-1}(i)) / \hat{\rho}_{p-1} \\ \hat{a}_p(0) = 1 \\ \hat{a}_p(n) = \hat{a}_{p-1}(n) + \hat{K}_p \cdot \hat{a}_{p-1}(p-n); n = 2, \dots, p-1 \\ \hat{a}_p(n) = \hat{K}_p \\ \hat{\rho}_p = \hat{\rho}_{p-1} (1 - \hat{K}_p^2) \end{array} \right. \quad 2.31$$

Dans l'algorithme de Burg on part de l'expression de l'erreur de prédiction « avant » au rang p :

$$\hat{\varepsilon}_p^f(i) = \sum_{n=0}^p \hat{a}_p(n) \cdot x(i-n) \quad 2.32$$

et de l'erreur de prédiction rétrograde :

$$\hat{\varepsilon}_p^b(i) = \sum_{n=0}^p \hat{a}_p(p-n) \cdot x(i-n) \quad 2.33$$

Pour trouver un estimateur de K_p , Burg propose de minimiser, par rapport à K_p , la somme des erreurs de prédiction avant et rétrograde $\frac{1}{2}(\hat{\rho}_p^f + \hat{\rho}_p^b)$ avec :

$$\hat{\rho}_p^f = \frac{1}{N-p} \cdot \sum_{n=p}^{N-1} \left| \hat{\varepsilon}_p^f(n) \right|^2 \quad 2.34$$

$$\hat{\rho}_p^b = \frac{1}{N-p} \cdot \sum_{n=0}^{N-1-p} \left| \hat{\varepsilon}_p^b(n) \right|^2 \quad 2.35$$

La minimisation de ce terme par rapport au coefficient de réflexion donne [kay88], [kunt91]:

$$\hat{K}_p = \frac{-2 \sum_{n=p}^{N-1} \hat{\varepsilon}_{p-1}^f(n) \cdot \hat{\varepsilon}_{p-1}^b(n)}{\sum_{n=p}^{N-1} ((\hat{\varepsilon}_{p-1}^f(n))^2 + (\hat{\varepsilon}_{p-1}^b(n))^2)} \quad 2.36$$

donc on peut calculer les coefficients de réflexion d'une manière récursive sur l'ordre. En même temps on peut calculer les paramètres AR en utilisant (2.31) et en connaissant la relation :

$$\begin{aligned}
\hat{\varepsilon}_p^f(n) &= \hat{\varepsilon}_{p-1}^f(n-1) + \hat{K}_p \cdot \hat{\varepsilon}_p^b(n-1) \\
\hat{\varepsilon}_p^b(n) &= \hat{\varepsilon}_{p-1}^b(n-1) + \hat{K}_p \cdot \hat{\varepsilon}_p^f(n-1)
\end{aligned}
\tag{2.37}$$

Une propriété de l'algorithme de Burg est de garantir que les estimées des coefficients de réflexion sont de module inférieur à 1.
En résumé l'algorithme de Burg s'écrit comme suit :

$$\begin{aligned}
\bullet \hat{R}_{xx}(0) &= \frac{1}{N} \cdot \sum_{n=0}^{N-1} (x(n))^2 \\
\bullet \hat{\rho}_0 &= \hat{R}_{xx}(0) \\
\bullet \hat{\varepsilon}_0^f(n) &= x(n), n = 1, 2, \dots, N-1 \\
\bullet \hat{\varepsilon}_0^b(n) &= x(n), n = 1, 2, \dots, N-1 \\
\bullet \text{Pour } k &= 1, \dots, p \\
* \hat{K}_k &= \frac{-2 \sum_{n=k}^{N-1} \hat{\varepsilon}_{k-1}^f(n) \cdot \hat{\varepsilon}_{k-1}^b(n)}{\sum_{n=k}^{N-1} ((\hat{\varepsilon}_{k-1}^f(n))^2 + (\hat{\varepsilon}_{k-1}^b(n))^2)} \\
* \hat{\rho}_k &= (1 - \hat{K}_k^2) \cdot \hat{\rho}_{k-1} \\
* \hat{a}_k(i) &= \begin{cases} \hat{a}_{k-1}(i) + \hat{K}_k \cdot \hat{a}_{k-1}(k-i), & \text{pour } i = 1, 2, \dots, k-1 \\ \hat{K}_k, & \text{pour } i = k \end{cases} \\
* \hat{\varepsilon}_k^f(n) &= \hat{\varepsilon}_{k-1}^f(n-1) + \hat{K}_k \cdot \hat{\varepsilon}_k^b(n-1), \text{ pour } n = k+1, k+2, \dots, N-1 \\
* \hat{\varepsilon}_k^b(n) &= \hat{\varepsilon}_{k-1}^b(n-1) + \hat{K}_k \cdot \hat{\varepsilon}_k^f(n-1), \text{ pour } n = k, k+1, \dots, N-2
\end{aligned}
\tag{2.38}$$

Méthode de Covariance

La méthode de Yule-Walker minimise l'erreur de prédiction sur l'ensemble des données $x(n)$, ce qui oblige à compléter la séquence par des zéros au début et à la fin des données.

Dans la méthode de covariance on fait la minimisation uniquement sur les données. Il faut donc minimiser l'équation (2.28) et on arrive alors aux équations [kay88] :

$$\begin{pmatrix} C_{xx}(1,1) & C_{xx}(1,2) & \dots & C_{xx}(1,p) \\ C_{xx}(2,1) & C_{xx}(2,2) & \dots & C_{xx}(2,p) \\ \cdot & \cdot & \cdot & \cdot \\ C_{xx}(p,1) & C_{xx}(p,2) & \dots & C_{xx}(p,p) \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \cdot \\ \hat{a}_p \end{pmatrix} = - \begin{pmatrix} C_{xx}(1,0) \\ C_{xx}(2,0) \\ \cdot \\ C_{xx}(p,0) \end{pmatrix} \quad 2.39$$

et la variance de bruit d'entrée est :

$$\hat{\sigma}_\varepsilon^2 = \min_{n=p}^{N-1} (\sum \varepsilon_p^{f^2}(n)) = C_{xx}(0,0) + \sum_{k=1}^p \hat{a}(k).C_{xx}(0,k) \quad 2.40$$

avec

$$C_{xx}(j,k) = \frac{1}{N-p} \sum_{n=p}^{N-1} x(n-j).x(n-k); 0 \leq j \leq p \text{ et } 0 \leq k \leq p \quad 2.41$$

On note que la matrice de covariance n'est pas de Toeplitz et que l'algorithme de Levinson ne peut pas être utilisé pour la résolution du système linéaire. La matrice est par contre symétrique définie positive, donc l'algorithme de Cholesky peut être utilisé pour la résolution du système [marp87].

On note aussi que le système calculé par la méthode de covariance ne garantit pas la stabilité [kay88]. Cependant cette instabilité partielle est rarement rencontrée en pratique [kunt91].

Méthode de Covariance modifiée

Cette méthode utilise la même démarche que précédemment mais en minimisant cette fois la somme des erreurs de prédiction avant et arrière.

On minimise donc la somme :

$$\sum_{n=p}^{N-1} \varepsilon_p^{f^2}(n) + \sum_{n=0}^{N-1-p} \varepsilon_p^{b^2}(n) \quad 2.42$$

On arrive aux mêmes formes d'équations que pour la méthode de covariance mais avec :

$$C_{xx}(j,k) = \frac{1}{2.(N-p)} (\sum_{n=p}^{N-1} x(n-j).x(n-k) + \sum_{n=0}^{N-1-p} x(n+i).x(n+p)) \quad 2.43$$

Problème de récursivité

Dans la littérature, des algorithmes récursifs de type LMS et RLS [marp87] sont décrits pour la première fenêtre (fenêtres « before »). Ces algorithmes sont cependant lourds à utiliser car leur principe repose sur une méthode itérative, pour laquelle se posent des problèmes de choix des conditions initiales et des paramètres de convergence.

En général, pour les algorithmes d'estimation AR (dans les cas de Yule-Walker, de covariance et de covariance modifiée), le temps de calcul le plus important concerne le calcul d'autocorrélation ou de covariance, et non l'inversion de matrices d'autocorrélation ou

de covariance (car ce sont des matrices d'ordre p généralement faible). Il est donc utile de rechercher la récursivité, essentiellement pour les calculs d'autocorrélation et de covariance.

Pour l'algorithme de Burg, il est très difficile d'avoir une récursivité sur les échantillons (car le principe même de l'algorithme utilise tous les échantillons disponibles).

Pour l'algorithme de Yule-Walker, on peut voir facilement que pour la fenêtre allant de 1 à j - p :

$$\hat{R}_{xx}^{(j)}(k) = \frac{1}{j-p} \sum_{n=1}^{j-p-k} x(n).x(n+k) \quad 2.44$$

alors :

$$(j-p).\hat{R}_{xx}^{(j)}(k) = (j-p-1).\hat{R}_{xx}^{(j-1)}(k) + x(j-p-k).x(j-p) \quad 2.45$$

et donc :

$$\hat{R}_{xx}^{(j)}(k) = \frac{j-p-1}{j-p} \hat{R}_{xx}^{(j-1)}(k) + \frac{1}{j-p} x(j-p-k).x(j-p) \quad 2.46$$

Pour la deuxième fenêtre de j+1 à j+N :

$$\hat{R}_{xx}^{(j)}(k) = \frac{1}{N} \sum_{n=j+1}^{j+N} x(n).x(n+k) \quad 2.47$$

donc à l'instant j+1 on a :

$$\begin{aligned} \hat{R}_{xx}^{(j+1)}(k) &= \frac{1}{N} \sum_{n=j+2}^{j+N+1} x(n).x(n+k) = \\ & \hat{R}_{xx}^{(j)}(k) + \frac{1}{N} [x(j+N+1).x(j+N+1+k) - x(j+1).x(j+1+k)] \end{aligned} \quad 2.48$$

On peut faire le même calcul pour les covariances dans les deux algorithmes de covariance et de covariance modifiée.

On voit donc qu'il est aisé de rendre récursifs (plus précisément semi-récursifs) les trois algorithmes d'estimation AR (Yule-Walker, covariance et covariance modifiée), de façon à réduire notablement le temps de calcul.

2.3.2 Utilisation de la modélisation AR

Dans notre étude un modèle AR a été utilisé pour estimer les coefficients AR (a_i) des deux types des signaux d'EMG : EMG d'accouchement prématuré (AP), EMG accouchement à terme (figure 2.2).

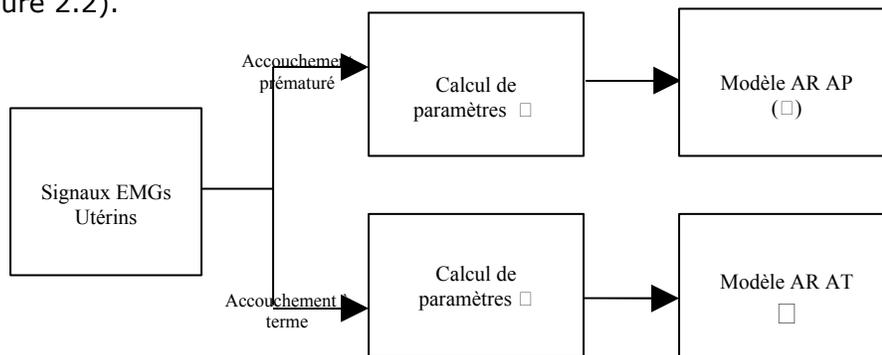


Figure 2.2. Création des modèles AR pour AP et AT.

Après obtention des paramètres AR des deux groupes, on génère des signaux qui simulent les deux types des contractions utérines : signaux EMGs d'accouchements prématurés simulés et signaux EMGs à termes simulés, en utilisant à l'entrée des modèles un bruit blanc gaussien.

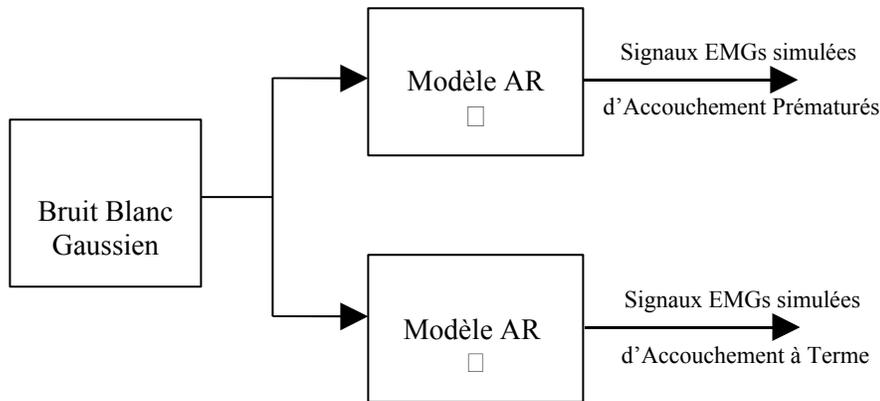


Figure 2.3. Génération des signaux EMGs AP simulés et signaux EMGs AT.

2.4 Estimation spectrale pour l'EMG

2.4.1 Estimation non paramétrique

Le périodogramme est une méthode d'estimation spectrale non paramétrique qui n'impose aucun modèle pour le signal. Elle est fondée sur la transformation de Fourier discrète.

Soit N échantillons du signal x_0, x_1, \dots, x_{N-1} , le périodogramme est défini par [marp87], [kay88], [orfa90], [kunt91] :

$$\hat{S}_{PER}(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x_n \cdot e^{-j2\pi fn} \right|^2 \quad 2.49$$

Il faut noter que cet estimateur est biaisé. De plus on montre que $Var(\hat{S}_{PER}(f)) \approx \left| \hat{S}_{PER}(f) \right|^2$

quel que soit N [kay88] [kunt91]. Cet estimateur donc est inconsistant, la variance du périodogramme étant très grande. Pour réduire cette variance Bartlett [bart48] [proa96] a proposé une modification au périodogramme. Dans la méthode de Bartlett les segments de taille N sont divisés en K segments de M points non superposés. Pour chaque segment on calcule le périodogramme, puis on définit le périodogramme de Bartlett comme étant la moyenne de ces périodogrammes élémentaires. Le périodogramme de Bartlett, appelé aussi périodogramme moyenné, diminue la variance par un facteur K , mais réduit en même temps la résolution spectrale du même facteur K . Il faut donc trouver un compromis dans le choix de K entre la résolution spectrale et la diminution de variance.

En réalité, parmi les méthodes d'estimation non paramétriques, on utilise plutôt la méthode du périodogramme moyenné modifié ou méthode de Welch [proa96]. Cette méthode diffère de la méthode de Bartlett par deux points essentiels : on applique une fenêtre non rectangulaire (fenêtre de Hanning en général) aux données, pour diminuer l'effet des lobes dans l'estimation des points spectraux. On utilise également des segments partiellement superposés. Pour chaque segment on définit le découpage de la façon suivante :

$$\begin{aligned} x_i(n) &= x(n+iD) \quad n=0,1,\dots,M-1 \\ i &= 0,1,\dots,L-1 \end{aligned} \quad 2.50$$

où M est la taille des segments après découpage, d est le décalage d'un segment à l'autre et L est le nombre de segments obtenus. Dans le cas d'un recouvrement de 50%, on a D=M/2 et on a alors L=2K segments. Le périodogramme modifié de chaque segment i est estimé par :

$$\hat{S}^i(f) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x_i(n) \cdot w(n) \cdot e^{-j2\pi fn} \right|^2 \quad 2.51$$

avec U facteur de normalisation de la fenêtre W, défini par :

$$U = \frac{1}{M} \sum_{n=0}^{M-1} w^2(n) \quad 2.52$$

L'estimation du spectre par la méthode de Welsh est alors définie par :

$$\hat{S}_W(f) = \frac{1}{L} \sum_{i=0}^{L-1} \hat{S}^i(f) \quad 2.53$$

La méthode de Welsh est la méthode non paramétrique la plus utilisée pour estimer un spectre [marp87], [kay88], [proa96], car elle réduit davantage la variance de l'estimateur que celle de Bartlett [marp87][orfa90].

2.4.2 Estimation paramétrique

On se place ici dans le cas de signaux modélisés par un modèle paramétrique de type AR (voir paragraphe 2.3). Les estimateurs spectraux paramétriques sont caractérisés par le fait qu'ils sont des estimateurs de résolution infinie. Ces estimateurs sont efficaces à condition que le modèle choisi soit très proche du signal réel, ce qui est très difficile à démontrer en pratique.

Dans le cas d'un processus AR, le spectre est obtenu par :

$$S_{AR}(f) = \frac{\sigma_\varepsilon^2}{\left| 1 + \sum_{k=1}^p a_k \cdot e^{-j2\pi fk} \right|^2} \quad 2.54$$

Dans le cas réel on doit estimer les paramètres des modèles à partir des données réelles, donc utiliser l'une des méthodes déjà discutées pour l'estimation d'un modèle AR. Dans le cas d'un modèle AR estimé, l'expression du spectre devient alors :

$$\hat{S}_{AR}(f) = \frac{\hat{\sigma}_\varepsilon^2}{\left| 1 + \sum_{k=1}^p \hat{a}_k \cdot e^{-j2\pi fk} \right|^2} \quad 2.55$$

Les paramètres du modèle sont estimés par l'une des méthodes d'estimation déjà citée dans le paragraphe 2.3. Dans la littérature, les avantages et les inconvénients de chaque

méthode vis à vis d'une estimation spectrale ont été largement décrits [marp87], [kay88], [proa96]. Ce paragraphe n'en fera donc qu'un bref résumé. Il est démontré que, de toutes les méthodes, la méthode d'autocorrélation donne le spectre le moins bon (on peut voir par exemple l'étude de simulation faite par Kay [kay88]). Ceci est dû à l'introduction de zéros dans la série de données, ce qui induit une mauvaise estimation des paramètres AR. La méthode de Burg est meilleure dans la mesure où elle présente une résolution infinie et produit un modèle AR stable [marp87] [proa96]. Cependant un inconvénient de la méthode est qu'elle fait apparaître des pics supplémentaires lorsque l'ordre du modèle est élevé [marp87], [kay88]. De plus elle peut aboutir à un décalage de fréquence dans certains cas [marp87], [kay88]. Les méthodes de Covariance et de Covariance modifiée sont également meilleures que celle d'auto-corrélation et résolvent le problème de décalage de fréquence [marp87], [kay88].

Le point le plus important dans le cas d'estimation paramétrique reste la recherche de l'ordre du modèle utilisé. Cependant, dans l'approche de classification, nous ne sommes pas tenus de choisir l'ordre du modèle le plus proche de la réalité, mais seulement un ordre adapté à la discrimination. Dans le cas d'estimation spectrale le problème se pose différemment car dans ce cas, un ordre trop faible fournit un spectre trop lissé, alors qu'un ordre très élevé produit des pics erronés [marp87], [kay88].

2.4.3 Estimation pseudo- spectrale par la méthode MUSIC

La méthode MUSIC n'est pas à proprement parler un estimateur spectral mais on peut la qualifier d'estimateur pseudo- spectral. Elle est attractive pour notre problème dans la mesure où elle permet une estimation dans le cas d'un rapport signal sur bruit défavorable. Cette méthode repose sur la décomposition de la matrice d'autocorrélation en deux sous espaces, espace du signal et espace du bruit [marp87], [kay88], [orfa90].

La méthode MUSIC a été utilisée dans différents domaines, par exemple dans la localisation de sources [redd79], la résolution d'échos [bruc85], ou encore dans le traitement de signaux biomédicaux [akay96].

Elle permet d'extraire des pics de fréquence, donc des sinusoides, noyées dans un bruit blanc additif [marp87], [geor00]. Elle nous a donc paru un moyen adapté de caractériser des signaux EMG bruités mais utilisables.

Le principe de la méthode MUSIC repose sur la décomposition de la matrice d'auto-corrélation en valeurs propres, les unes correspondant au signal et les autres au bruit. Supposons qu'un enregistrement contienne P sinusoides complexes noyées dans un bruit blanc additif de moyenne nulle et de variance σ_0^2 non corrélé avec le signal.

Le signal d'observation est de la forme :

$$y(n) = x(n) + b(n) \quad 2.56$$

$x(n)$ étant le signal composé de P sinusoides complexes qu'on peut écrire alors sous forme :

$$x(n) = A_1 e^{jw_1 n} + A_2 e^{jw_2 n} + \dots + A_p e^{jw_p n} = x_1(n) + x_2(n) + \dots + x_p(n) \quad 2.59$$

A_i est l'amplitude de la composante i du signal,

w_i est la fréquence de la composante i du signal.

On démontre [ther92] que chaque composante x_i du signal correspond à un vecteur propre de la matrice d'auto- corrélation R_y de valeur propre $N.E[|A_i|^2] + \sigma_0^2$.

Les (N-P) vecteurs propres restants sont orthogonaux aux composantes x_i et ont des valeurs propres égale à σ_0^2 .

Donc les (N-P) vecteurs propres qui correspondent aux valeurs propres de valeur σ_0^2 correspondent à ce qu'on appelle « l'espace du bruit », orthogonal à l'espace formé par les composantes \mathbf{x}_i qui correspondent à ce qu'on appelle « l'espace du signal ».

Reposant sur ce principe, plusieurs méthodes ont été développées pour estimer le pseudo spectre d'un tel signal. Parmi ces méthodes on peut citer la méthode de Pisarenko [marp87] qui estime le pseudo-spectre d'un signal formé de P sinusoides dans un échantillon de données de longueur (P+1).

La méthode MUSIC est une généralisation de la méthode de Pisarenko [marp87], [kay88], [akay96], [ther92]. L'algorithme est le suivant :

1- Calculer une estimation de la matrice d'auto-corrélation par :

$$\hat{R}_y = \begin{pmatrix} \hat{R}_y(0) & \hat{R}_y(1) & \dots & \hat{R}_y(N) \\ \hat{R}_y(1) & \hat{R}_y(0) & \dots & \hat{R}_y(N-1) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \hat{R}_y(N) & \hat{R}_y(N-1) & \dots & \hat{R}_y(0) \end{pmatrix} \quad 2.58$$

avec :

$$\hat{R}_y(k) = \frac{1}{N-k} \sum_{n=0}^{N-1-k} y(n+k).y(n) \quad 2.59$$

2- Calculer les valeurs propres et les vecteurs propres correspondants de la matrice d'auto-corrélation \hat{R}_y .

3- Calculer la dimension K de l'espace du bruit, égale au nombre des faibles valeurs propres, par un critère d'Akaike ou MDL (Minimum Description Length) [akay96], [ther92]. On en déduit la dimension de l'espace du signal égale à (N+1-K).

4- Calculer le pseudo- spectre MUSIC par :

$$\hat{S}(f) = \frac{1}{\frac{1}{k} \sum_{i=1}^{K-1} |E_i(w)|^2} \quad 2.60$$

les $E_i(w)$ étant les vecteurs propres de l'espace du bruit.

Le pseudo-spectre MUSIC ne peut pas être utilisé pour l'estimation des paramètres (moyenne, médiane....) du spectre EMG [falo02].

2.4.4 Discussion sur le choix de la méthode d'estimation du spectre des contractions

L'estimation de la densité spectrale de puissance d'un signal EMG est très importante car c'est à partir de cette densité que nous allons calculer les paramètres pertinents afin de les utiliser dans l'algorithme de classification.

L'estimation du spectre EMG par la méthode de Welch est la plus utilisée en pratique. Plus de 90% des auteurs utilisent cette méthode pour estimer le spectre EMG [duch93].

L'estimation du spectre EMG par des méthodes paramétriques, surtout par modèle AR, est parfois utilisée également pour estimer un spectre EMG [duch93], [merl99a], [merl99b]. Cependant le problème reste le choix de l'ordre du modèle, certains auteurs utilisant un modèle d'ordre 20 [Inba84], d'autres 15 [Seid87] pour des EMG de muscles striés. L'utilisation d'un modèle AR pour calculer le spectre EMG conduit à une bonne estimation spectrale. Il est bien connu qu'un estimateur de type AR est très sensible au bruit [kay88], ce qui a pour effet de diminuer la résolution spectrale. Mais nos signaux sont filtrés et présentent un bon rapport signal sur bruit.

Une comparaison de l'estimation spectrale par le périodogramme moyenné et par modélisation AR pour des applications EMG a été faite dans [merl99b]. Bien qu'il existe des différences dans l'estimation entre les deux méthodes pour les calculs des paramètres globaux (médiane, moyenne ...), les deux méthodes donnent des résultats similaires, et les valeurs de ces paramètres globaux varient très peu avec l'ordre du modèle [koro99]. La méthode paramétrique est préférée pour les petits segments (<0.25 s).

Nous avons donc dans notre étude utilisé le périodogramme de Welch qui produit une estimation beaucoup plus robuste que le modèle AR lorsque le niveau de bruit est important [kunt91], [kay88], ainsi que l'estimation par modèle AR.

2.4.5 Extraction des paramètres à partir de la densité spectrale

Bien que les paramètres AR d'un signal puissent être utilisés pour la comparaison, et la classification, nous avons utilisé ces paramètres seulement pour augmenter la taille de notre base de données. Ensuite Les paramètres de classifications ont été extraits directement de la densité spectrale des signaux réels ou simulés.

Puissance de signal

Elle représente la répartition de l'énergie M_0 (Moment d'ordre 0) du signal sur l'axe des fréquences, avec le Moment d'ordre r donné par l'expression :

$$M_r = 2 \int_0^{\infty} f^r S_x(f) df \quad 2.61$$

$S_x(f)$ est la densité spectrale du signal.

Fréquence moyenne

Elle représente la moyenne statistique du signal.

$$MPF = M_1 / M_0 \quad 2.62$$

Avec

$$M_r = 2 \int_0^{\infty} f^r S_x(f) df \quad 2.61$$

Coefficient de dissymétrie

Ce paramètre donne des informations concernant la forme de la densité spectrale du point de vue symétrique. Il est défini par:

$$CD = \frac{M_3^*}{\sqrt{M_2^{3*}}} \quad 2.63$$

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

Coefficient d'aplatissement (Kurtosis)

Le kurtosis caractérise la forme de pic ou l'aplatissement relatif d'une distribution comparée à une distribution normale. Un kurtosis positif indique une distribution relativement pointue, tandis qu'un kurtosis négatif signale une distribution relativement aplatie.

Il est donné par la formule :

$$CA = \frac{M_4^*}{M_2^{2*}} \quad 2.64$$

Fréquence médiane

La médiane partage la distribution de la densité spectrale en deux parties : 50% des données sont plus petites que la médiane, 50% sont plus grandes.

La médiane est calculée par:

$$\int_0^{F_{med}} S_x(f) df = \int_{F_{med}}^{F_{max}} S_x(f) df \quad 2.65$$

Energie relative par bande de fréquence / Déciles

On a vu que la médiane partage la distribution de la densité spectrale en deux parties. On peut généraliser le partage de cette distribution en quatre, en dix, en cent, ou en un nombre quelconque de parties. Les valeurs ainsi obtenues sont appelées des quartiles, des déciles, des percentiles (ou centiles), ou des quantiles. L'énergie de chaque intervalle est donnée par la formule :

$$W_n = \frac{\int_{f_{n-1}}^{f_n} S_x(f) df}{M_0} \quad 2.66$$

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

avec N le nombre d'intervalles.

Ces paramètres représentent la variance spectrale. L'axe de fréquence est distribué en dix intervalles égaux,

$$\int_{f_{r-1}}^{f_r} S_x(f) df = k \int_0^{F_{max}} S_x(f) df \quad 2.67$$

$$0 < k \leq 1$$

Fréquence pic

La Fréquence pic est la fréquence pour laquelle la [fonction de densité spectrale](#) atteint un maximum.

Entropie spectrale

L'entropie mesure la quantité d'information moyenne contenue dans un signal, elle est significative de la variance spectrale :

$$H = -\int_0^{f_{\max}} S_x(f) \ln[S_x(f)] df \quad 2.68$$

Rapport H/B (haut/bas)

Il s'agit d'un rapport entre deux bandes d'énergie extraites après études a priori des formes de la densité spectrale dans les situations d'intérêt (par exemple étude de la fatigue ...).

En général, une bande augmente quand l'autre diminue ou inversement. Cela donne une bonne différence et donc un bon paramètre de discrimination.

$$\frac{H}{L} = \frac{\int_{f_{H1}}^{f_{H2}} S_x(f) df}{\int_{f_{L1}}^{f_{L2}} S_x(f) df} \quad 2.69$$

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

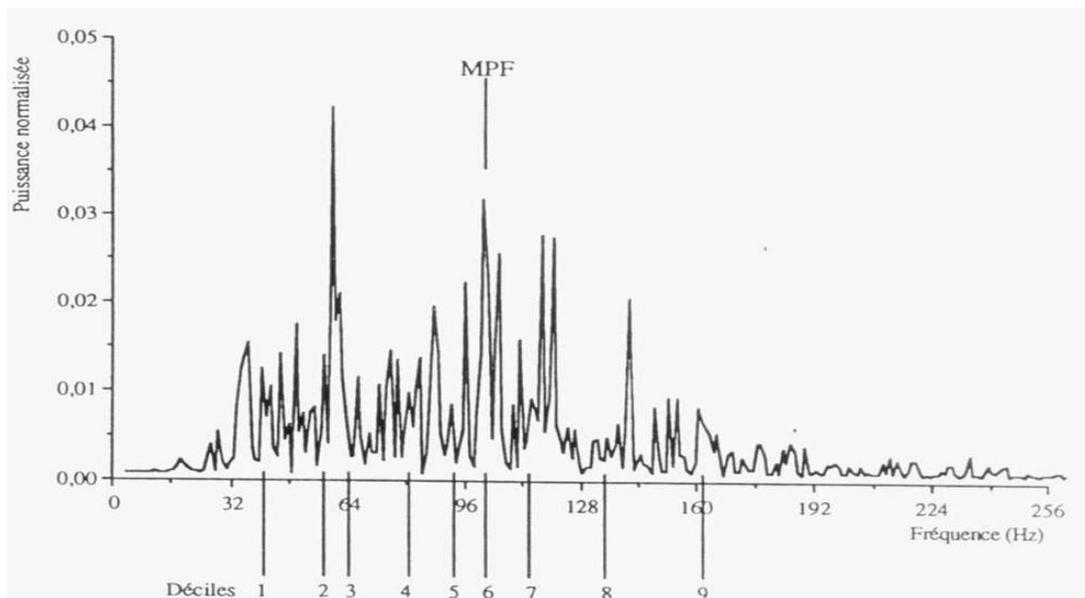


Figure 2.4 : Paramétrage du spectre [hogr94].

La figure 2.4 figure présente les déciles qui permettent de témoigner de la forme fine du spectre d'un EMG, la MPF représente sa moyenne statistique, l'entropie est significative de la variance spectrale et ne peut être représentée sur cette figure.

2.5. Conclusion

La décomposition en ondelettes peut être utilisée pour extraire des paramètres pertinents. Après avoir fait la décomposition en coefficients de détails, les variances sur les échelles sont utilisées comme paramètres pour la classification.

La modélisation par modèle AR a été utilisée dans deux buts : tout d'abord afin de générer une grande base de données à partir de signaux EMG utérin. Cette partie est discutée ultérieurement. Ces paramètres ensuite extraits de cette densité spectrale de puissance sont les déciles.

La modélisation spectrale reposant sur l'estimation par le périodogramme de Welch a été retenue pour sa capacité à produire une estimation spectrale dont il est possible de réduire l'influence d'un bruit blanc additif. C'est donc cette représentation qui sera utilisée ensuite afin d'effectuer une extraction de paramètres significatifs.

Chapitre 3

Méthodes de classification

3.1 Introduction

Ce chapitre donne une présentation sur les différentes méthodes de classification. Après avoir calculé les paramètres des contractions, deux types de classification peuvent être appliqués :

- *classification non supervisée* : on ne prend pas en compte les classes connues a priori des contractions. Dans ce cas, la méthode de classification consiste à faire un « Clustering » des nuages des points dans l'espace. Ensuite une interprétation est faite pour affecter chaque cluster à une classe physiologique. Ces méthodes sont avantageuses quand nous n'avons pas un ensemble d'apprentissage assez grand pour faire un apprentissage correct pour certaines méthodes de classification supervisées.

- *classification supervisée* : on prend en compte les classes de contractions connues a priori pour l'apprentissage. Nous avons utilisé une méthode basée sur la modélisation AR pour augmenter d'abord l'ensemble d'apprentissage et rendre aussi l'apprentissage possible.

Nous commencerons par décrire les méthodes de classifications supervisées tout d'abord avec les méthodes paramétriques, comme le rapport de vraisemblance. Nous présenterons ensuite les méthodes non paramétriques comme la méthode du K plus proche voisin et la méthode de Parzen.

Ensuite nous décrivons les réseaux de neurones artificiels, les perceptrons multicouches et les réseaux à fonctions radiales. Nous discuterons ensuite leurs architectures, des équations de fonctionnement et de la stratégie d'apprentissage.

La dernière partie de ce chapitre est dédiée à la présentation des différentes méthodes de classification non supervisées. Nous commencerons par la description de la méthode K moyenne, la méthode de LBG et la méthode de K-moyenne flou. Puis nous décrivons les réseaux compétitifs et leurs apprentissages compétitifs pour présenter enfin une méthode originale de classification: méthode de classification non supervisée basée sur le test de Fisher (USCM, unsupervised statistical classification method).

La technique d'analyse en composante principale est présentée ensuite à la fin de ce chapitre. Cette technique a été utilisée pour projeter les données dans un espace des paramètres réduits (généralement R^2) pour faciliter l'interprétation et la visualisation.

3.2 Classification supervisée

Dans le cas où la distribution des paramètres est connue, on peut utiliser des méthodes de classification paramétrique basées sur le rapport de vraisemblance. Dans notre cas, la distribution des paramètres est inconnue. Nous avons donc utilisé des méthodes de classification non paramétriques comme les méthodes de KPPV et Parzen. De plus les réseaux de neurones peuvent être utilisés dans tous les cas.

3.2.1 Méthode paramétrique : rapport de vraisemblance

Le rapport de vraisemblance est utilisé dans le cas où les paramètres des classes sont connus. Définissons :

- $P(x/w_1)$: la probabilité de distribution des vecteurs sous la classe 1.
- $P(x/w_2)$: la probabilité de distribution des vecteurs sous la classe 2.
- $P(w_i)$: les probabilités a priori des classes / $\sum P(w_i) = 1$.
- $P(w_i/x)$: la probabilité de classier le vecteur dans la classe 1.

x est classifié dans w_1 si :

$$P(w_1/x) > P(w_2/x) \quad 3.1$$

x est classifié dans w_2 si :

$$P(w_2/x) > P(w_1/x) \quad 3.2$$

Mais

$$P(w_1/x) = \frac{P(x/w_1)P(w_1)}{P(x)} \quad 3.3$$

avec

$$P(x) = P(x/w_1).P(w_1) + P(x/w_2).P(w_2) \quad 3.4$$

x est classifié dans w_1 si :

$$\frac{P(x/w_1)P(w_1)}{P(x)} > \frac{P(x/w_2).P(w_2)}{P(x)} \quad 3.5$$

ou

$$\frac{P(x/w_2)}{P(x/w_1)} < \frac{P(w_1)}{P(w_2)} \quad 3.6$$

on peut démontrer aussi que x est classifié dans w_2 si :

$$\frac{P(x/w_1)}{P(x/w_2)} < \frac{P(w_2)}{P(w_1)} \quad 3.7$$

Appelons le rapport de vraisemblance :

$$n(x) = \frac{P(x/w_2)}{P(x/w_1)} \quad 3.8$$

donc on peut écrire le test de classification :

$$n(x) = \frac{P(x/w_2) < P(w_1)}{P(x/w_1) > P(w_2)} \quad 3.9$$

Considérons par exemple le cas gaussien :

$$\begin{aligned} x \in \mathfrak{R}^n & \quad x \propto N(\mu_1, \Sigma_1) \text{ dans } w_1 \\ & \quad x \propto N(\mu_2, \Sigma_2) \text{ dans } w_2 \end{aligned} \quad 3.10$$

$$P(x/w_1) = \frac{1}{(2\pi)^{N/2} |\Sigma_1|} \text{Exp}^{-\frac{1}{2}(x-\mu_1)^T \Sigma_1^{-1} (x-\mu_1)} \quad 3.11$$

Donc le test de classification devient :

$$-\frac{1}{2}(x - \mu_2)\Sigma_2^{-1}(x - \mu_2) + \frac{1}{2}(x - \mu_1)\Sigma_1^{-1}(x - \mu_1) \underset{>}{\leq} \text{Log} \frac{P(w_1)}{P(w_2)} + 0.5 \text{Log} \frac{|\Sigma_2|}{|\Sigma_1|} \quad 3.12$$

3.2.2 Méthode non paramétrique : Règle du K plus proche voisin

La méthode du plus proche voisin se caractérise par sa capacité à traiter localement les données. Elle consiste à examiner les éléments, dont la classe est connue, proches de l'élément dont on veut déterminer la classe.

La règle du plus proche voisin, PPV, compte sur l'existence d'un ensemble test. Soit $X^n = \{\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_n\}$ l'ensemble d'apprentissage formé de n vecteurs indépendants, où chaque événement est représenté par un vecteur paramétrique, \mathbf{x}_j , ayant k éléments. 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 des éléments de l'ensemble d'apprentissage. La règle dite de 1PPV est simplement la suivante :

$$\hat{w}(\mathbf{x}) = w(\mathbf{x}_{PPV}) \quad \text{si} \quad d(\mathbf{x}, \mathbf{x}_{PPV}) = \underset{j=1 \dots n}{\text{Min}} d(\mathbf{x}, \mathbf{x}_j) \quad 3.13$$

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 de 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 [gova03].

3.2.3 Méthode non paramétrique : Parzen

L'approche de Parzen a été proposée par Rosenblatt (1956) [rose56] dans le cas unidimensionnel puis par Parzen (1962). Cette méthode est basée sur l'estimation de la densité de probabilité dans un volume donné.

Supposons que la région R_n est un hypercube de côté h_n et de dimension d, n est le nombre des points totaux à utiliser. Le volume est :

$$V(R_n) = V_n = h_n^d \quad 3.14$$

Soit la fonction :

$$\phi(u) = \begin{cases} 1 & \text{si } |u^j| < 1/2 \quad \forall j \in \{1, \dots, d\} \\ 0 & \text{sinon} \end{cases} \quad 3.15$$

Le nombre de points qui tombent dans l'hypercube centré sur x est:

$$k_n = \sum_{i=1}^n \phi\left(\frac{x - x_i}{h_n}\right) \quad 3.16$$

Dans ce cas l'estimateur de la densité au point x devient:

$$p_n(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{V_n} \phi\left(\frac{x - x_i}{h_n}\right) \quad 3.17$$

Pour garantir que $p_n(x)$ soit une densité, il faut avoir les conditions suivantes :

$$\phi(u) \geq 0 \quad 3.18$$

et

$$\int \phi(u) du = 1 \text{ (avec } V_n = h_n^d) \quad 3.19$$

Le choix du paramètre d'échelle a une influence importante sur $p(x)$

$$\delta_n(x) = \frac{1}{V_n} \phi\left(\frac{x}{h_n}\right) \quad 3.20$$

$$P_n(x) = \frac{1}{n} \sum_{i=1}^n \delta_n(x - x_i) \quad 3.21$$

Si h_n est trop grand, la densité estimée ne sera pas précise du tout.

Si h_n est petit, cette dernière sera considérablement variable. Il est possible de donner à chaque classe une valeur différente de h_n .

3.2.4 Les réseaux de neurones

Introduction

Les réseaux de neurones sont des processeurs distribués parallèles qui possèdent la propriété de mémoriser des expériences passées en vue de les utiliser dans des procédés non connus à l'avance. Ils ressemblent au cerveau par deux aspects :

- La connaissance est acquise par le réseau au moyen d'un processus d'apprentissage.
- Les intensités des connexions entre les neurones connues par les 'poids synaptiques' sont utilisées pour mémoriser la connaissance.

La procédure utilisée pour effectuer le processus d'apprentissage est appelée algorithme d'apprentissage. Cet algorithme a pour rôle de modifier les poids synaptique du réseau afin d'obtenir un but désiré (classification par exemple).

De plus, il est possible qu'un réseau de neurones modifie sa propre topologie. Cette modification se trouve motivée par le fait que les neurones du cerveau peuvent mourir et d'autres connexions synaptiques peuvent se créer en fonction des besoins.

D'après ce qui précède, il apparait que les réseaux de neurones trouvent leur puissance dans :

- Leur structure qui est à distribution parallèle.
- Leur pouvoir d'apprendre pour ensuite généraliser.

Issus de travaux à connotation biologique dans les années 1940, ces réseaux sont maintenant considérés comme des outils mathématiques, indépendamment de toute référence à la biologie. Ils sont utilisés pour la modélisation et la commande de processus non linéaires, ainsi que comme outils de classification, notamment pour la reconnaissance de formes.

Les principales étapes dans l'évolution de la théorie et de la pratique des réseaux de neurones ont été la mise au point d'un algorithme d'apprentissage, économique en temps de calcul, pour l'évaluation du gradient de la fonction de coût, appelé *algorithme de rétro-propagation* [rume86], et la preuve de ses propriétés d'approximateur universel [horn89] et de parcimonie [barr93, Horn94]. L'une des premières applications dans le domaine de la modélisation non linéaire de processus est présentée dans [nare90].

Réseaux de Neurones : Perceptron multicouches

Les réseaux de neurones possèdent la propriété d'approcher 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 "rétro-propagation de l'erreur" qui se base sur la règle d'apprentissage par correction d'erreur. Le processus de rétro-propagation 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 se répètent 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.

Architecture du réseau

La figure 3.1 montre l'architecture d'un Perceptron Multicouches avec (L) couches ayant (n) entrées et (m) sorties. Le réseau est entièrement connecté, c.a.d. chaque neurone d'une couche est relié à l'ensemble des neurones de la couche suivante.

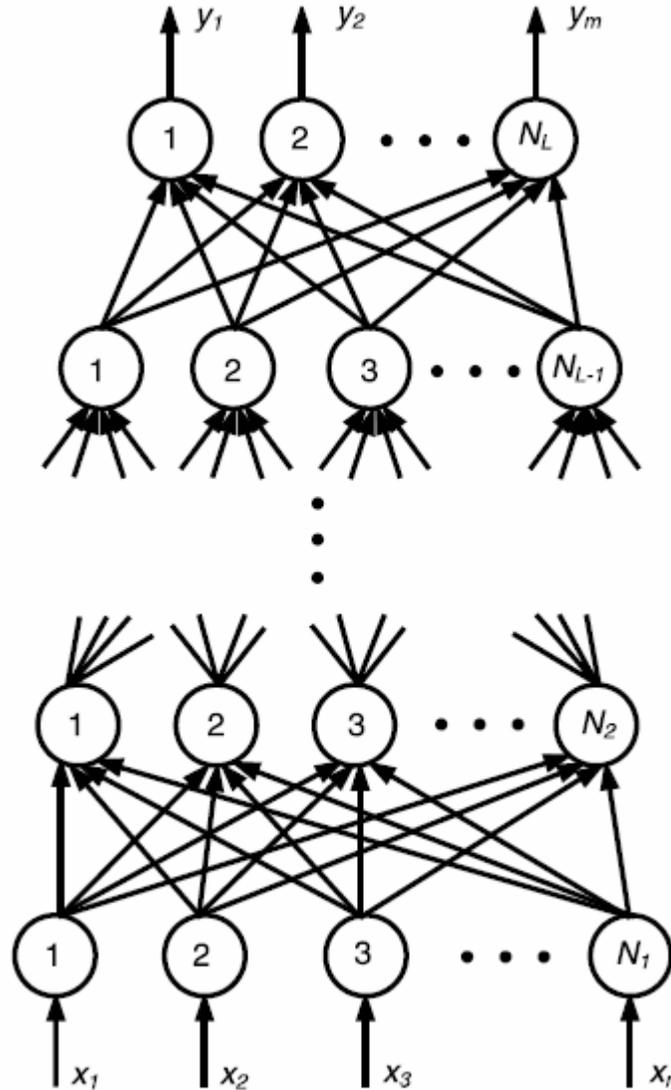


Figure 3.1. Structure d'un réseau MLP. Le nombre total des couches est L . La 1^{ère} couche est la couche d'entrée, la $L^{\text{ième}}$ couche est la couche de sortie, et les couches de 2 à $L-1$ représentent les couches cachées [zhan00].

Traitement de l'information par un neurone

Dans un réseau de neurone, chaque neurone (excepté les neurones de la couche d'entrée) reçoit et traite des stimuli (entrées) d'autres neurones. L'information traitée est disponible à la sortie du neurone. Le schéma 3.2 illustre la manière dont chaque neurone dans un MLP traite l'information.

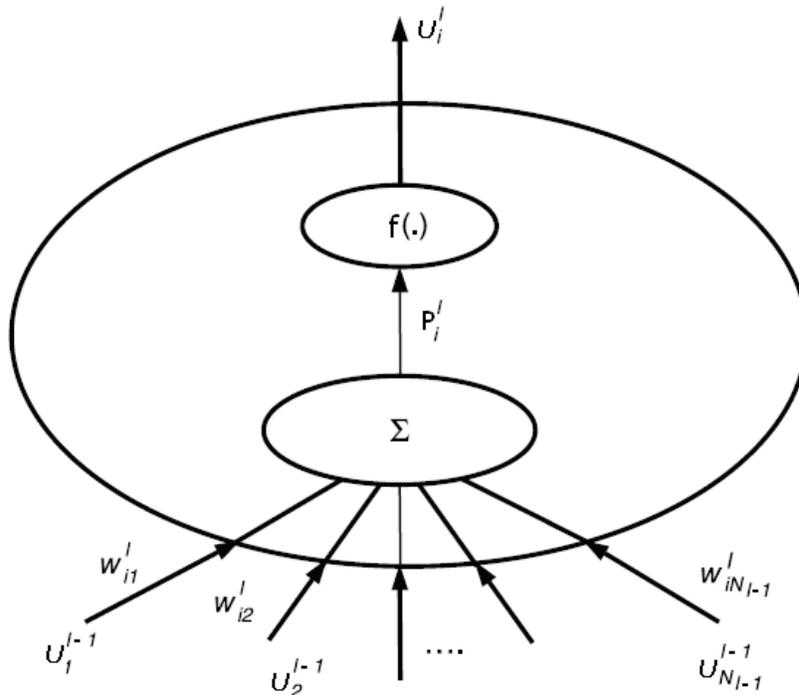


Figure 3.2. Exemple d'un neurone .

Comme exemple, un neurone de la $l^{i\text{ème}}$ couche reçoit des stimuli des neurones de la couche $l-1$, c.-à-d., $U_1^{l-1}, U_2^{l-1}, \dots, U_{N_{l-1}}^{l-1}$. Chaque entrée est d'abord multipliée par le paramètre de poids correspondant, et les produits résultants sont ajoutés pour produire une somme P . Cette somme est passée par une fonction d'activation de neurone $f(\cdot)$ pour produire la sortie finale du neurone. La sortie U_i^l , à son tour, devient un stimulus pour les neurones de la prochaine couche.

Quelques fonctions d'activation

Différentes fonctions de transfert peuvent être utilisées comme fonction d'activation du neurone. Les trois les plus utilisées sont les fonctions «seuil» «linéaire» et «sigmoïde».

La fonction seuil

Cette fonction est définie par :

$$f(\gamma) = \begin{cases} 1 & \text{si } \gamma > 0 \\ 0 & \text{si } \gamma < 0 \end{cases} \quad 3.22$$

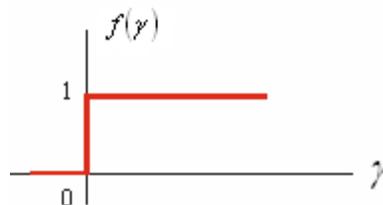


Figure 3.3. Fonction seuil.

Cette fonction est utilisée dans les réseaux de type perceptrons.

La fonction linéaire

Cette fonction est définie par :

$$f(\gamma) = \gamma \quad 3.23$$

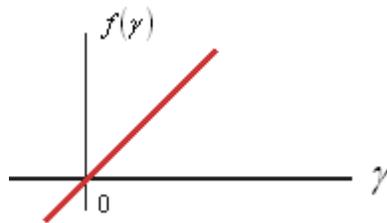


Figure 3.4. Fonction linéaire.

Cette fonction est utilisée avec les approximateurs linéaires.

La fonction sigmoïde

Cette fonction est définie par :

$$f(\gamma) = \frac{1}{1 + \text{Exp}^{-a\gamma}} \quad 3.24$$

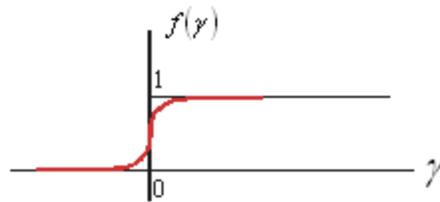


Figure 3.5. Fonction sigmoïde.

Cette fonction est la plus utilisée car elle est différentiable avec :

$$\frac{d\varphi(\gamma)}{d\gamma} = a(\varphi(\gamma) - \varphi^2(\gamma)) \quad 3.25$$

Algorithme de retro-propagation

La rétro-propagation est actuellement l'outil le plus utilisé dans le domaine des réseaux de neurones. C'est une technique de calcul des dérivées qui peut être appliquée à n'importe quelle structure de fonctions dérivables.

Mathématiquement, cette méthode est basée sur l'algorithme de descente du gradient et utilise les règles de dérivation des fonctions dérivables. Dans cette méthode, l'erreur commise en sortie du réseau sera rétropropagée vers les couches cachées d'où le nom de rétro-propagation.

Equations du réseau:

Avant de définir la règle d'apprentissage, on doit définir la relation entre les sorties du réseau d'une part, et les entrées et les poids d'autre part.

Dans un réseau à (l) couches ayant (n) entrées et (m) sorties les états des différents neurones sont donnés par :

$$U_i^l(k) = f^l(P_i^l(k)) \quad 3.26$$

avec

$$P_i^l(k) = \sum_{j=1}^{N_{l-1}} W_{ij}^l U_j^{l-1}(k) \quad 3.27$$

Ou $i=1,2,\dots,N_l$
 et $j=1,2,\dots,N_{l-1}$
 et $q=1,2,\dots,N_{l+1}$

N_l nombre de neurones dans la couche l.
 N_{l-1} nombre de neurones dans la couche l-1.
 N_{l+1} nombre de neurones dans la couche l+1.
 L nombre de couches.

$U_i^1(k)$ sortie du neurone i de la couche 1.

$W_{ij}^1(k)$ coefficient synaptique (poids) de la j^{eme} entrée et les sorties du réseau.

$$U_i^0(k) = X_i(k) \quad i=1,2,\dots,n$$

$$U_i^1(k) = Y_i(k) \quad i=1,2,\dots,m$$

Ou : $X_i(k)$ et $Y_i(k)$ sont respectivement les entrées et les sorties du réseau.

L'objectif de la méthode de rétro-propagation est d'adapter les paramètres W_{ij}^1 de façon à minimiser une fonction de coût donnée par :

$$E(W) = \sum_{p=1}^T E_p(W) \quad 3.28$$

avec

$$E_p(W) = \frac{1}{2} \sum_{i=1}^m [y_i^d(k) - y_i(k)]^2 \quad 3.29$$

Ou $y^d(k)$ est le vecteur de sortie désiré, $y(k)$ le vecteur de sortie du réseau et T le nombre d'exemples, ou longueur de l'ensemble d'apprentissage.

L'approche la plus utilisée pour la minimisation de la fonction E est basée sur la méthode du gradient. On commence l'apprentissage par un choix aléatoire des vecteurs initiaux de poids. On présente le premier vecteur d'entrée. Une fois obtenue à la sortie du réseau l'erreur correspondante, les gradients de l'erreur par rapport à tous les poids sont calculés. Les poids sont alors ajustés. On refait la même procédure pour tous les exemples d'apprentissage. Ce processus est répété jusqu'à ce que les sorties du réseau soient suffisamment proches des sorties désirées.

Adaptation des poids

L'adaptation des poids se fait par la méthode du gradient basée sur la formule itérative suivante :

$$W_{ij}^l(k+1) = W_{ij}^l(k) - \Delta W_{ij}^l \quad 3.30$$

avec

$$\Delta W_{ij}^l = l_r \frac{\partial E(W)}{\partial W_{ij}^l(k)} \quad 3.31$$

Où k : représente le numéro d'itération.

l_r est une constante appelée facteur ou pas d'apprentissage.

La vitesse de convergence dépend de la constante l_r . Sa valeur est généralement choisie expérimentalement.

La dérivée de la fonction du coût par rapport au poids W_{ij}^l est donnée par :

$$\frac{\partial E(w)}{\partial W_{ij}^l(k)} = \sum_{p=1}^T \frac{\partial E_p(W)}{\partial W_{ij}^l(k)} \quad 3.32$$

où T est le nombre des points d'apprentissage.

$$\frac{\partial E(w)}{\partial W_{ij}^l(k)} = \frac{\partial E_p(W)}{\partial U_i^l(k)} \cdot \frac{\partial U_i^l(k)}{\partial W_{ij}^l(k)} \quad 3.33$$

pour la couche de sortie :

$$\frac{\partial E_p(w)}{\partial U_i^l(k)} = -(y_i^d(k) - y_i(k)) \quad 3.34$$

pour les couches cachées :

$$\frac{\partial E_p(w)}{\partial U_i^l(k)} = \sum_{q=1}^{N_{l+1}} \frac{\partial E_p(w)}{\partial U_q^{l+1}(k)} \cdot \frac{\partial U_q^{l+1}(k)}{\partial U_i^l(k)} \quad 3.35$$

$$\frac{\partial U_i^l(k)}{\partial W_{ij}^l} = f^l(P_i^l(k)) \cdot U_j^{l-1}(k) \quad 3.36$$

$$\frac{\partial E_p(W)}{\partial W_{ij}^l} = \frac{\partial E_p(W)}{\partial U_j^l(k)} \cdot f^l(P_j^l(k)) \cdot U_j^{l-1}(k) \quad 3.37$$

Pour minimiser l'erreur totale sur l'ensemble d'apprentissage, les poids du réseau doivent être ajustés après présentation de tous les exemples.

Algorithme

Etape 1 : initialiser les poids W_{ij}^l et les seuils internes des neurones à des faibles valeurs aléatoires.

Etape 2 : présenter le vecteur d'entrée et la sortie désirée, correspondant.

Etape 3 : Calculer la sortie du réseau en utilisant les expressions (3.26) et (3.27)

Etape 4 : Calculer l'erreur de sortie en utilisant l'expression (3.34)

Etape 5 : Calculer l'erreur dans les couches en utilisant l'expression (3.35)

Etape 6 : Calculer le gradient de l'erreur par rapport aux poids en utilisant l'expression (3.31)

Etape 7 : Ajuster les poids selon l'expression (3.30).

Etape 8 : Si la condition sur l'erreur ou sur le nombre d'itérations est atteinte, aller à l'étape 9, sinon aller à l'étape 2.

Etape 9: Fin.

Les exemples sont présentés d'une manière successive. Lorsque tous les exemples ont été présentés, le test s'effectue sur l'erreur de sortie et les poids sont ajustés au fur et à mesure, jusqu'à ce que l'erreur de sortie se stabilise à une valeur acceptable.

Les réseaux à fonctions radiales (RBF pour Radial Basis Functions)

Les réseaux RBF forment une autre famille de réseaux de neurones. On va tout d'abord voir comment est constitué un tel réseau, puis ensuite décrire leur apprentissage.

Architecture du réseau

Le réseau RBF (Radial Basis Functions) est un réseau à deux couches [pogg89]. Les neurones de la première couche sont reliés aux entrées et ont chacun deux paramètres : un vecteur prototype P et un coefficient d'étalement σ strictement positif. La fonction d'activation réalisée, de forme gaussienne, est la suivante :

$$f_1(X) = e^{-\frac{\|X-P\|^2}{2\sigma^2}} \quad 3.38$$

Le vecteur prototype P définit un point dans l'espace d'entrée. La sortie s du neurone est égale à 1 pour une entrée x égale à P , puis décroît vers 0 lorsque l'entrée s'éloigne de P . La vitesse de décroissance est réglée par σ : plus le coefficient est petit et plus la fonction sera concentrée autour du point P et proche de 0 ailleurs.

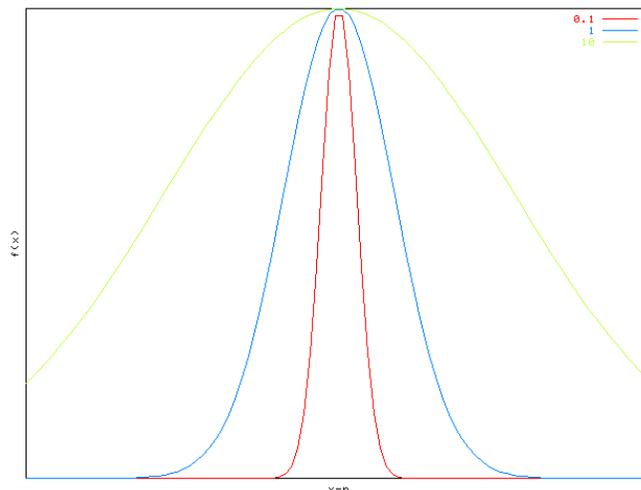


Figure 3.6. Gaussiennes obtenues avec différents étalements ($\sigma=0,1$ 1 et 10) [tert03].

Les neurones de la seconde couche quand à eux calculent la sortie du réseau en effectuant une combinaison linéaire des sorties de ceux de la première couche, avec un biais qui est ajouté au total. La fonction qu'ils réalisent est la suivante :

$$f_2(X) = V \cdot W + b \quad 3.39$$

où x est un vecteur composé des sorties de tous les neurones de la première couche, w est un vecteur de poids et b est le biais. W et b sont ajustés lors de l'apprentissage. Ainsi le réseau peut être représenté par le schéma suivant :

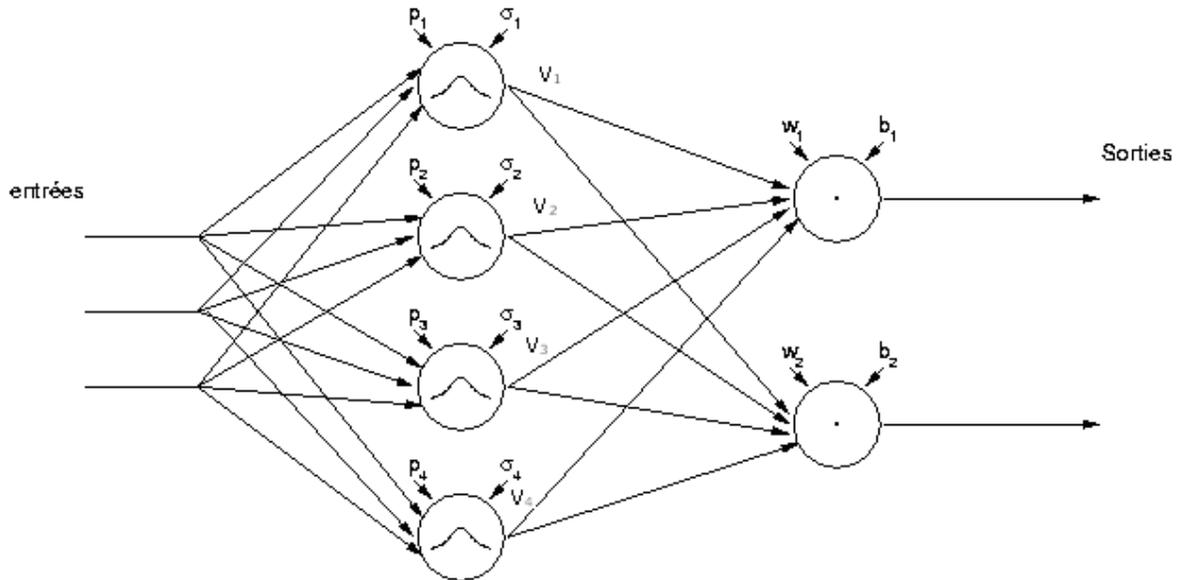


Figure 3.7. Schéma général d'un réseau RBF [tert03].

Chaque sortie du réseau RBF est donnée par la formule :

$$s_i(x) = \sum_{j=1}^{N_g} w_{i,j} e^{-\frac{\|x-p_j\|^2}{2\sigma_j^2}} + b_i \quad 3.40$$

où i est le numéro de la sortie (c'est à dire le numéro du neurone de la seconde couche dont on calcule la sortie), N_g le nombre de neurones de la première couche, P_j le vecteur prototype du neurone numéro j de la première couche, σ_j son coefficient d'étalement, $w_{i,j}, j = 1 \dots N_g$ les N_g poids du neurone de sortie i , et b_i son biais.

Il a été montré que le réseau RBF est un approximateur universel [park91], c'est à dire que le réseau est capable d'approximer n'importe quelle fonction douce avec une précision donnée, pourvu que l'on fournisse un nombre suffisant de neurones, et que l'on utilise un algorithme d'apprentissage adéquat. Lors de l'apprentissage d'un réseau RBF deux problèmes se posent : la constitution de la première couche (choix du nombre de neurones, choix des prototypes et des coefficients d'étalement), et la détermination des poids et biais de la seconde couche.

Apprentissage de la seconde couche

Une fois les prototypes placés et les coefficients d'étalement, il reste à déterminer les paramètres de la seconde couche, c'est-à-dire les poids w_i et biais b_i qui minimisent la fonction de performance du réseau. On peut calculer la sortie de la première couche pour chaque exemple de la base d'apprentissage :

$$Z_{k,j} = e^{-\frac{\|x_k - p_j\|^2}{2\sigma_j^2}} \quad 3.41$$

où k est le numéro de l'exemple et j celui de la composante de Z_k calculée, c'est-à-dire le numéro du neurone de la première couche dont on calcule la sortie. Le critère à minimiser est l'erreur quadratique moyenne du réseau sur la base d'apprentissage et a donc pour expression :

$$e = \frac{1}{N_a} \sum_{k=1}^{N_a} \sum_{i=1}^{N_s} (s_i(x_k) - y_{k,i})^2 = \frac{1}{N_a} \sum_{k=1}^{N_a} \sum_{i=1}^{N_s} (z_k \cdot w_i + b_i - y_{k,i})^2 \quad 3.42$$

Où $s_i(x_k)$ est la sortie du neurone i de la seconde couche (équation 3.40), $y_{k,i}$ est la i^{me} composante de y_k (c'est à dire la sortie voulue pour l'exemple k sur le neurone i de la seconde couche) et N_s est le nombre de sorties. Un algorithme des moindres carrés peut être appliqué afin de trouver les paramètres w_i et b_i qui minimisent ce critère.

Cette méthode d'apprentissage, qui consiste d'abord à sélectionner des prototypes puis ensuite à optimiser les poids de la seconde couche est simple à mettre en œuvre. D'autres algorithmes, tels que EM (Expectation-Maximisation) peuvent être appliqués aux RBF afin de déterminer ces deux jeux paramètres en même temps, avec souvent une meilleure performance du réseau, mais aux prix d'un temps d'apprentissage plus long [bish95]. Cependant toutes ces méthodes nécessitent de savoir par avance le nombre de prototypes à définir dans la première couche. D'autres méthodes ont donc été étudiées, de type incrémental, c'est à dire que des prototypes sont ajoutés dans la première couche au fur et à mesure des besoins de l'apprentissage.

3.3 Classification non supervisée

3.3.1 Méthode de K-moyenne

L'algorithme des K-moyennes est un algorithme classique de quantification vectorielle. Son principe est le suivant: « on dispose de points de l'espace des observations que l'on souhaite rassembler en classes, sans que l'on dispose de connaissance *a priori* de propriété(s) particulière(s) sur ces classes ; seul leur nombre p est fixé *a priori* » [drey02].

Algorithme

- 1) définir K centroides (vecteurs) initialisés au hasard.
- 2) « attribuer » chaque point au centroïde le plus proche (distance).
- 3) Calculer les coordonnées de chaque centroïde comme la moyenne des points qui lui sont attribués.
- 4) Tant que les centroides évoluent, aller en 2.

C'est un algorithme qui converge. Il est très utilisé en traitement du signal pour réduire le nombre de couleurs dans une image ou pour la compression par exemple [sapo90].

3.3.2 Méthode de Linde-Buzo-Gray (LBG)

L'algorithme de LBG ne fait pas seulement évoluer les centroïdes mais aussi leur nombre, par duplication [lind80]. L'intérêt de le faire est d'accroître la précision et aussi d'équilibrer les tailles des classes.

Algorithme LBG

Entrée : K centroïdes $D = \{y_1, \dots, y_K\}$, centres de gravité des K ensembles de la partition de $X = \{x_j, j = 1, \dots, n\}$. Un vecteur z.

Tant que critère d'arrêt non-satisfait Répéter

Doubler le nombre de centroïdes en remplaçant chaque centroïde y par deux autres y+z et y-z.

Affecter chaque exemple x_j à la classe du nouveau centroïde dont il est le plus proche.

Adapter les centroïdes en les remplaçant par les centres de gravité de leurs classes.

Sortie : un regroupement en K' classes et les centroïdes de chaque classe (leurs centres de gravité). L'avantage de cette méthode est que le nombre de classe n'est pas fixé a priori.

3.3.3 Méthode de K-moyenne flou

L'algorithme des K-moyennes floues (Fuzzy C-Means, FCM) est un algorithme de classification floue fondé sur l'optimisation d'un critère quadratique de classification où chaque classe est représentée par son centre de gravité. Cette méthode nécessite de connaître le nombre de classes au préalable et génère les classes par un processus itératif en minimisant une fonction objectif.

L'algorithme utilise l'ensemble des paramètres $A = \{x_1, x_2, \dots, x_n\}$ où x_i est un vecteur à plusieurs composantes, et le nombre de régions est égale à c. Les valeurs des degrés d'appartenance sont regroupées dans une matrice $U = [u_{ik}]$ pour : $1 \leq i \leq n$, $1 \leq k \leq c$, où u_{ik} désigne le degré d'appartenance du segment i à la classe k. Pour avoir une bonne partition, on impose aux éléments de U les contraintes suivantes :

$$u_{ik} \in [0,1]; \sum_k u_{ik} = 1; \forall i \quad 3.43$$

L'algorithme du FCM fait évoluer la partition (Matrice U) en minimisant la fonction suivante :

$$J_m(U, c) = \sum_i \sum_k (u_{ik})^m \cdot \|x_i - c_k\|^2 \quad 3.44$$

où : $m > 1$ est un paramètre contrôlant le degré de flou (généralement $m = 2$) ; c_k : le centre de la classe k.

3.3.4 Réseaux de neurones compétitifs

Le réseau de neurones compétitif (Competitive Neural Network Method, CNNM) permet aux neurones de sortie de se concurrencer afin qu'un seul neurone soit activé à un instant donné.

Pour le neurone gagnant, le niveau de son activité interne pour un signal d'entrée doit être le plus grand par rapport aux autres neurones du réseau. Le signal de sortie du neurone gagnant sera égal à un alors que les signaux de sortie des neurones perdants seront égaux à zéro.

Par conséquent, chaque neurone d'un réseau apprend à se spécialiser sur un ensemble de signaux d'entrée et devient un détecteur spécialisé.

Dans la forme la plus simple, le réseau compétitif possède un réseau de neurones à une seule couche de neurones de sortie, chacun étant complètement connecté aux signaux d'entrée.

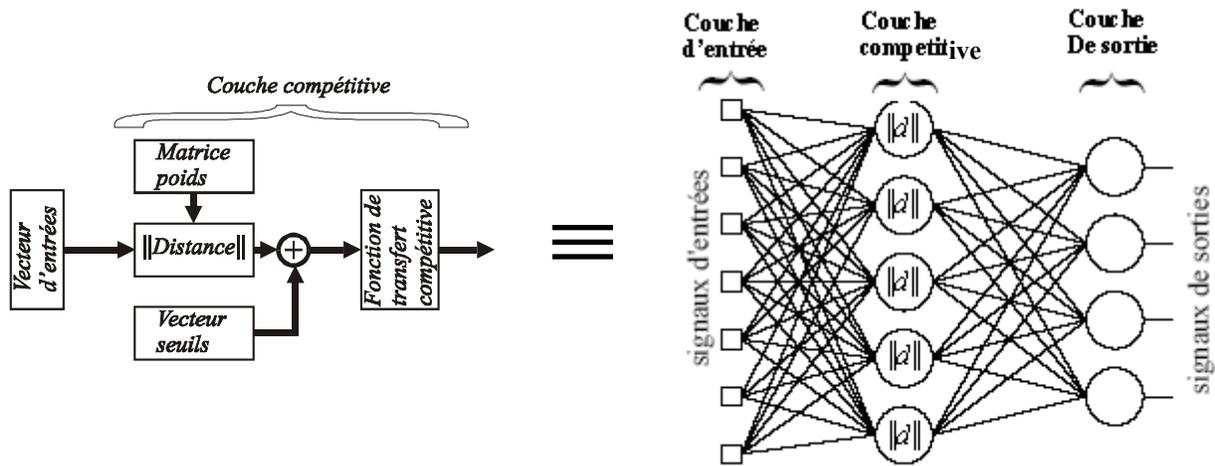


Figure 3.8. Réseau compétitif [must03].

Le réseau peut inclure des connexions latérales de poids fixes entre les neurones permettant d'effectuer des inhibitions latérales telles que chaque neurone tend à inhiber le neurone auquel il est relié.

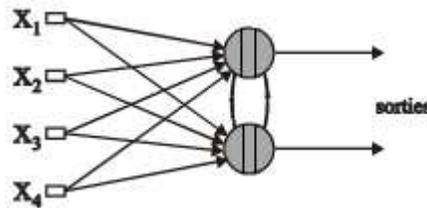


Figure 3.9. Connexions latérales [must03].

Si un neurone particulier gagne la compétition, alors chaque entrée de ce neurone gagnant change son poids tout en gardant une somme unité. L'ajustement appliqué au poids synaptique sera :

$$\Delta w_{kj} = \begin{cases} \eta(x_j - w_{kj}) & \text{si le neurone } k \text{ est gagnant} \\ 0 & \text{si le neurone } k \text{ est perdant} \end{cases} \quad 3.45$$

Dans ce cas, chaque neurone sera le spécialiste d'un groupe d'échantillons se présentant à l'entrée du réseau.

La loi de modification des poids des connexions est dérivée de celle de Hebb [hebb49]. A la présentation d'un vecteur d'entrée, un neurone de sortie est sélectionné. Il correspond le plus possible à cette entrée (minimisation d'une distance).

Apprentissage compétitif

Trois éléments de base existent pour l'apprentissage compétitif :

1. Un ensemble de neurones semblables, excepté pour quelques poids synaptiques aléatoires, peut répondre différemment à un ensemble de signaux d'entrée.
2. Une limite est imposée à la sortie de chaque neurone.
3. Un mécanisme existe qui permet aux neurones de se concurrencer, afin de répondre à un sous-ensemble des entrées tel que seulement un neurone sera activé à un instant donné. Le neurone vainqueur sera appelé neurone gagnant.

Dans le cas où les exemples d'entrées sont des vecteurs à deux composantes, l'algorithme d'apprentissage sera le suivant :

1. Initialisation des poids à des valeurs aléatoires autour d'une valeur centrale.
2. Présentation d'une entrée $X = (x_1, x_2)$.
3. Calcul de la distance de chacun des neurones par rapport à x_1 et x_2 .

$$d_j = |w_{j1} - x_1|^2 + |w_{j2} - x_2|^2 \quad 3.46$$

4. Sélection du neurone le plus proche $Min(d) = di$.
5. Modification des poids pour le neurone choisi (i). η est un paramètre correspondant au pas de modification des poids.

$$\begin{aligned} w_{i1} &= w_{i1} + \eta (x_1 - w_{i1}) \\ w_{i2} &= w_{i2} + \eta (x_2 - w_{i2}) \end{aligned} \quad 3.47$$

6. Tant que les performances sont insuffisantes : Retour à l'étape 2 et sélection de l'exemple suivant dans la base d'apprentissage [cich93].

3.3.5 Classification statistique basé sur la loi de Fisher-Snedecor et test de comparaison de variances

La loi de Fisher, liée au rapport de deux formes quadratiques indépendantes joue un grand rôle en statistique (loi du rapport des variances de deux échantillons indépendantes par exemple) [sapo90].

Si X et Y sont 2 variables aléatoires qui suivent les lois normales centrées réduites, alors :

$Z = \sum_{i=1}^n X_i^2$ et $T = \sum_{i=1}^p Y_i^2$, Z et T étant deux variables qui suivent indépendamment les lois

χ_n^2 et χ_p^2 (à n et p degrés de liberté), on définit :

$$F(n, p) = \frac{Z/n}{T/p} \quad 3.48$$

La densité de F s'obtient aisément par transformation de celle d'un bêta II car $X/2$ et $Y/2$ suivent des lois $\mathcal{Y}_{n/2}$ et $\mathcal{Y}_{p/2}$:

$$g(f) = \frac{1}{B\left(\frac{n}{2}, \frac{n}{2}\right)} \frac{\left(\frac{n}{p}\right)^{n/2} f^{n/2-1}}{\left(1 + \frac{n}{p}f\right)^{(n+p)/2}} \quad 3.49$$

$$E(F) = \frac{p}{p-2} \quad 3.50$$

et

$$V(F) = 2 \frac{p^2}{n} \frac{n+p-2}{n(p-2)^2(p-4)} \quad 3.51$$

Cette loi permet d'obtenir les distributions de la loi bêta I et bêta II ; on a en effet les relations suivantes :

-si Y suit une loi bêta II(n, p), alors $\frac{pY}{n}$ est un $F(2n, 2p)$;

-si X suit une loi bêta I(n, p), alors $\frac{p}{n} \frac{X}{1-X}$ est un $F(2n, 2p)$.

Tests de Fisher-Snedecor pour échantillons indépendants

Dans Fisher le problème se formalise de la manière suivante :

- $F_1(x)$: fonction de répartition d'un variable aléatoire X_1 .

- $F_2(x)$: fonction de répartition d'un variable aléatoire X_2 .

$$\begin{cases} H_0 : F_1(x) = F_2(x) \\ H_1 : F_1(x) \neq F_2(x) \end{cases} \quad 3.52$$

Le choix de H_0 est dicté par des considérations pratiques car $F_1(x) \neq F_2(x)$ est trop vague pour obtenir une région critique.

Dans la pratique on se contentera de vérifier l'égalité des espérances et des variances de X_1 et X_2 , en disposant de \bar{x}_1 et \bar{x}_2 et s_1^2 et s_2^2 moyennes et variances empiriques des deux échantillons si les lois de X_1 et X_2 sont gaussiennes [sapo90].

Dans le cas de deux échantillons gaussiens $X_1 \in LG(m_1, \sigma_1)$ et $X_2 \in LG(m_2, \sigma_2)$, les hypothèses deviennent alors :

$$H_0 : m_1 = m_2 \text{ et } \sigma_1 = \sigma_2 \quad \text{contre} \quad H_1 : m_1 \neq m_2 \text{ et } \sigma_1 \neq \sigma_2$$

Le test va consister à tester d'abord les variances, et si elles ne sont pas significativement différentes, à tester ensuite les espérances en admettant $\sigma_1 = \sigma_2$.

En appliquant les résultats de la théorie de l'échantillonnage :

$$\frac{n_1 S_1^2}{\sigma_1^2} \in \chi_{n_1-1}^2 \quad 3.53$$

et

$$\frac{n_2 S_2^2}{\sigma_2^2} \in \chi_{n_2-1}^2 \quad 3.54$$

Dans l'hypothèse $H_0 : \sigma_1 = \sigma_2$ on a :

$$F_{n_1-1, n_2-2} = \frac{\frac{n_1 S_1^2}{n_1 - 1}}{\frac{n_2 S_2^2}{n_2 - 1}} \quad 3.55$$

On peut interpréter F comme le rapport de deux estimateurs de σ_1^2 et σ_2^2 respectivement. Si $\sigma_1 = \sigma_2$, ce rapport ne doit pas différer significativement de 1. F sera la variable de décision. En pratique on met toujours au numérateur la plus grande des deux quantités :

$$\frac{n_1 S_1^2}{n_1 - 1} \quad \text{et} \quad \frac{n_2 S_2^2}{n_2 - 1}$$

Et la région critique est de la forme $F > k$ avec $k > 1$.

Si les deux échantillons ont la même taille $n_1 = n_2 = n$, le calcul se simplifie et :

$$F_{n-1, n-1} = \frac{S_1^2}{S_2^2} \quad 3.56$$

Si le test de Fisher-Snedecor aboutit à la conclusion $\sigma_1 = \sigma_2$, on passe au test des espérances.

3.4 Méthode originale : méthode de classification non supervisée basée sur le test de Fisher

Après extraction des paramètres à partir de la décomposition en détails d'ondelettes, chaque contraction est caractérisée par le vecteur des paramètres. La méthode de classification statistique non supervisée (USCM, unsupervised statistical classification method) consiste à trouver les vecteurs les plus proches pour les regrouper ensemble dans une même classe. Pour cela, les vecteurs sont comparés en utilisant le Test de Fisher mélangé avec la méthode de K-moyenne.

L'algorithme se résume de la façon suivante :

1. Initialiser le premier vecteur, comme centre de la première classe.
2. Pour chaque nouveau vecteur, calculer la vraisemblance entre ce vecteur et le centre des classes en utilisant le test de Fisher [sapo90] :

$$F_{n_1-1; n_2-2} = \frac{\frac{n_1 S_1^2}{n_1 - 1}}{\frac{n_2 S_2^2}{n_2 - 1}} \quad 3.57$$

n_1, n_2 sont les longueurs de deux signaux comparés.

S_1, S_2 sont les variances de deux signaux comparées (signaux des détails décomposés par ondelette).

En fait, nous avons utilisé le test de Fisher car on suppose que les échantillons du signal suivent une loi normale. Dans ce cas, les variances suivent une loi de Fisher.

3. Si les vecteurs (nouveau vecteur et centre de classe) sont identiques avec une probabilité de confiance donnée, adapter les classes. Sinon créer une nouvelle classe.
4. Retour à l'étape 2.

Un seuil est utilisé pour comparer les vecteurs au centre (intervalle de confiance de Fisher). Une grande valeur de seuil aboutit à un nombre minimal de classes. Si la valeur du seuil est petite, alors on obtient un grand nombre des classes.

3.5 Analyse en composantes principales (ACP)

L'analyse en composantes principales est une méthode vectorielle linéaire de réduction des dimensions de paramètres, en choisissant les directions dont la variance du nuage de données est la plus grande. Les données peuvent alors être visualisées sur un nombre de dimensions réduit (2 ou 3).

L'ACP se calcule à partir de la matrice de covariance des données. Celle-ci est diagonalisée afin d'en extraire les valeurs et vecteurs propres. Les données sont projetées dans l'espace défini par les vecteurs propres. Les valeurs propres, classées dans l'ordre décroissant, correspondent dans l'espace d'arrivée au vecteur propre dont la direction maximise la variance.

Toutefois, l'ACP n'est pas une méthode de classification ; si les nuages ont individuellement de grandes variances sur une dimension de l'ACP alors celle-ci sera jugée Composante Principale. Elle ne permet pas la séparation des classes.

Soit la matrice des données contenant m objets (ou individus) et p descripteurs (ou variables) :

x_{11}	·	·	x_{1p}	3.58
·	x_{ij}	x_{ij}	·	
·	x_{rj}	·	·	
x_{m1}	·	·	x_{mp}	
\bar{x}_1	\bar{x}_i	\bar{x}_r	\bar{x}_p	
s_1	s_i	s_r	s_p	

La centrage et réduction des données est fait par :

$$X_{ij} = \frac{x_{ij} - \bar{x}_i}{s_i} \quad 3.59$$

Alors les distances entre objets sont :

$$d(i, i')^2 = \sum_{j=1}^p p_j (X_{ij} - X_{i'j})^2 \quad 3.60$$

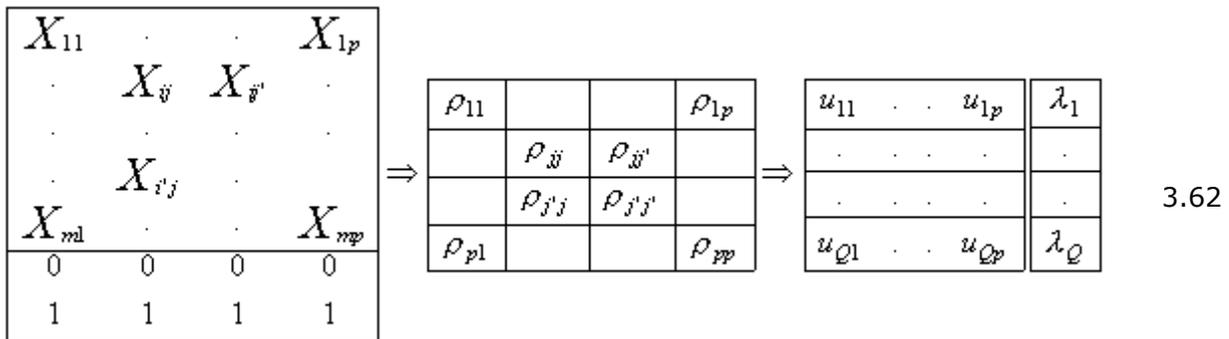
où p_j est une pondération (en général égale à 1).

L'inertie totale I du nuage est égale à :

$$I = \frac{1}{m} \sum_{j=1}^p \sum_{i=1}^m \left(\frac{x_{ij} - \bar{x}_j}{S_j} \right)^2 = p \quad 3.61$$

Le calcul des axes orthogonaux (indépendants) qui ont la propriété d'extraire le maximum de la variance projetée (ou inertie projetée) des individus ou objets, procède comme suit :

- calcul de la matrice des données centrées-réduites.
- calcul de la matrice des corrélations variables/variables.
- calcul de la matrice des vecteurs propres et du vecteur des valeurs propres.



Avec $Q \leq p$

- Les vecteurs propres sont les coordonnées du point de norme 1 des nouveaux axes sur les anciennes variables. Les valeurs sont proportionnelles à la variance associée à ces axes.
- Calcul des coordonnées des individus sur les nouveaux axes.
- Calcul des corrélations des variables avec les nouveaux axes.

La relations de transitions est définie par:

$$F_q(i) = \frac{1}{\sqrt{\lambda_q}} \sum_{j=1}^p \frac{x_{ij} - \bar{x}_j}{S_j} G_q(j) \quad 3.63$$

Où F est la projection de l'individu i sur l'axe factoriel q, G les projections des p variables sur q, et l la valeur propre associée à la composante q.

ET par réciprocity:

$$G_q(j) = \frac{1}{\sqrt{\lambda_q}} \sum_{i=1}^p \frac{x_{ij} - \bar{x}_j}{S_j} F_q(i) \quad 3.64$$

3.6 Matrice de confusion

Il s'agit d'un tableau de contingence confrontant les classes obtenues (colonnes) et les classes désirées (lignes). Sur la diagonale, on trouve donc les valeurs bien classées, hors de la diagonale on trouve les éléments mal classés. La somme des valeurs sur une ligne donne le nombre d'exemplaires théoriques de la catégorie.

Si les classes sont indépendantes, la position de l'erreur n'a aucune signification. Si par contre les classes ne sont pas indépendantes, on peut définir une sorte de gradation dans les erreurs.

Exemple:

Prenons un exemple de classification, 3 classes A, B et C indépendantes la matrice de confusion est la suivante:

	A	B	C
A	50	0	0
B	4	45	1
C	6	4	40

On voit que l'on disposait d'un échantillon de 150 valeurs réparties équitablement entre les 3 classes, ce qui n'est pas une nécessité. La classe A est correctement discriminée, les deux autres classes posent quelques difficultés (erreurs):

- 4 valeurs de B sont considérées dans A et une valeur dans C
- 6 valeurs de C sont considérées dans A et 4 valeurs dans B

Si l'on regarde les trois indicateurs:

- Taux global: on a un taux de reconnaissance de 90%(50+45+40/150=90), et un taux d'erreur de 10%.

3.7 Conclusion

Dans ce chapitre nous avons présenté les différentes méthodes de classification : les méthodes de classification supervisées et les méthodes de classification non-supervisées. Nous avons présenté la description des réseaux de neurones artificiels (MLP, RBF, CNNM), leur architecture, les équations de fonctionnement et la stratégie d'apprentissage.

Le contenu fréquentiel des contractions change d'une femme à une autre et durant la grossesse. Pour cela, nous avons proposé une méthode de classification non supervisée

originale basée sur la décomposition en ondelette et le test statistique et Fisher. La décomposition par ondelette est utilisée pour l'extraction des paramètres de chaque contraction et une méthode de classification non supervisée basée sur le test de Fisher est utilisée pour classier les évènements. Une projection par Analyse en Composantes Principales permet de visualiser les groupes résultants de la méthode de classification. Les résultats des cette méthodes vont être comparés ultérieurement avec les autres méthodes de classification traditionnelles. Le chapitre suivant va traiter une nouvelle méthode de classification basée sur les réseaux d'ondelettes.

Chapitre 4

Nouvelle approche : Réseau d'ondelette

4.1 Introduction

Dans le chapitre précédent, nous avons présenté les méthodes de classification qui existent dans la littérature, aussi qu'une nouvelle approche basée sur le test de Fisher. Nous avons vu aussi que les paramètres utilisés sont les coefficients d'ondelettes (les variances des détails). Ces paramètres sont utilisés comme entrée d'un réseau de neurones ou d'un algorithme de classification classique.

La nouvelle idée de ce chapitre vient du fait, qu'au lieu de faire la décomposition en ondelette puis de faire la classification par un réseau de neurones, nous allons utiliser directement un réseau d'ondelette, qui est un réseau de neurone standard mais avec des fonctions d'activation qui ressemblent aux ondelettes (dilatations et translations). Au lieu d'utiliser le signal, qui se présente comme une composante aléatoire, nous avons utilisé les déciles calculés à partir de la DSP. Ce type réseau qui a beaucoup été utilisé pour la régression est utilisé ici pour la classification.

4.2 Neurone formel à fonction dorsale

Pour un neurone formel à fonction dorsale, le calcul de la fonction est effectué en deux étapes :

1. Calcul d'une somme pondérée des entrées. Le résultat obtenu est appelé *potentiel* du neurone. Il est donné par la relation suivante :

$$v_i = \sum_{j \in P_i} c_{ij} x_j \quad 4.1$$

où P_i est l'ensemble des indices $\{j\}$ des entrées x_j du neurone i . Les coefficients c_{ij} sont des coefficients de pondération des entrées des neurones appelés poids synaptiques ou plus simplement poids.

2. Calcul d'une fonction non linéaire du potentiel (fonction sigmoïde), souvent appelée fonction d'activation. La sortie du neurone est alors la valeur de cette fonction, appelée parfois activité du neurone :

$$x_i = f(v_i) \quad 4.2$$

Ainsi, en tout point d'un hyperplan de l'espace des entrées défini par $v_i = \text{constante}$, la sortie du neurone est constante. Si le neurone est une fonction de deux variables, les lignes de niveau de la sortie sont des droites parallèles, d'où le terme de fonction dorsale.

La sortie d'un neurone à fonction dorsale est non linéaire par rapport aux entrées et par rapport aux coefficients c_{ij} .

Il est commode de représenter graphiquement un neurone à fonction dorsale comme indiqué sur la Figure 4.1, où apparaissent les deux phases du calcul de la sortie.

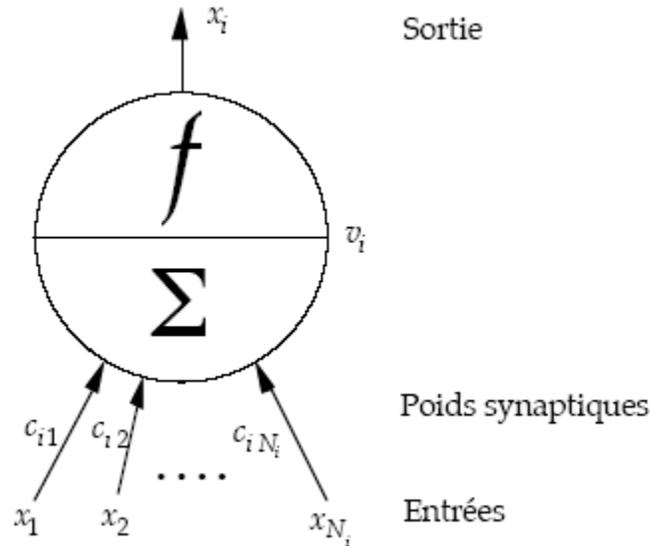


Figure 4.1. Représentation graphique d'un neurone à fonction dorsale [Ouss98].

4.3 Définition des ondelettes multidimensionnelles et des réseaux d'ondelettes

4.3.1 Ondelettes multidimensionnelles

Dans le chapitre 2, nous avons présenté les ondelettes à une dimension. Dans le cadre de la modélisation et de classification, il est fréquent d'avoir affaire à des processus multivariés, il est donc utile d'introduire la notion d'ondelette multidimensionnelle.

On peut définir une ondelette multidimensionnelle comme le produit d'ondelettes monodimensionnelles : on dit alors que les ondelettes sont séparables. Dans ce cas, l'expression d'une ondelette multidimensionnelle est :

$$\Phi_j(x) = \prod_{k=1}^{N_i} \phi(z_{jk}) \quad \text{avec} \quad z_{jk} = \frac{x_k \pm m_{jk}}{d_{jk}} \quad 4.3$$

où x_k est la k -ième composante du vecteur d'entrée x , et z_{jk} la composante centrée par m_{jk} et dilatée d'un facteur d_{jk} . Il a été montré dans [kuga95] que ces ondelettes multidimensionnelles sont des structures obliques de $L^2(\mathbf{R}^{N_i})$.

4.3.2 Réseaux d'ondelettes

Dans le présent travail, nous considérons des réseaux d'ondelettes de la forme suivante :

$$y = \psi(x) = \sum_{j=1}^{N_w} c_j \Phi_j(x) + \sum_{k=0}^{N_i} a_k x_k \quad \text{avec} \quad x_0 = 1 \quad 4.4$$

où y est la sortie du réseau et $x = \{ x_1, x_2, \dots, x_{N_i} \}$ le vecteur des entrées ; il est souvent utile de considérer, outre la décomposition en ondelettes proprement dite (première partie de l'équation), que la sortie peut avoir une composante affine par rapport aux variables, dont les coefficients sont les a_k ($k = 0, 1, \dots, N_i$). Pour la simplicité de l'exposé, nous ne considérerons que des réseaux à une sortie ; la généralisation à des réseaux à plusieurs sorties ne présente pas de difficulté.

Une ondelette peut être représentée de manière analogue à un neurone, comme indiqué sur la figure 4.2 [ouss98].

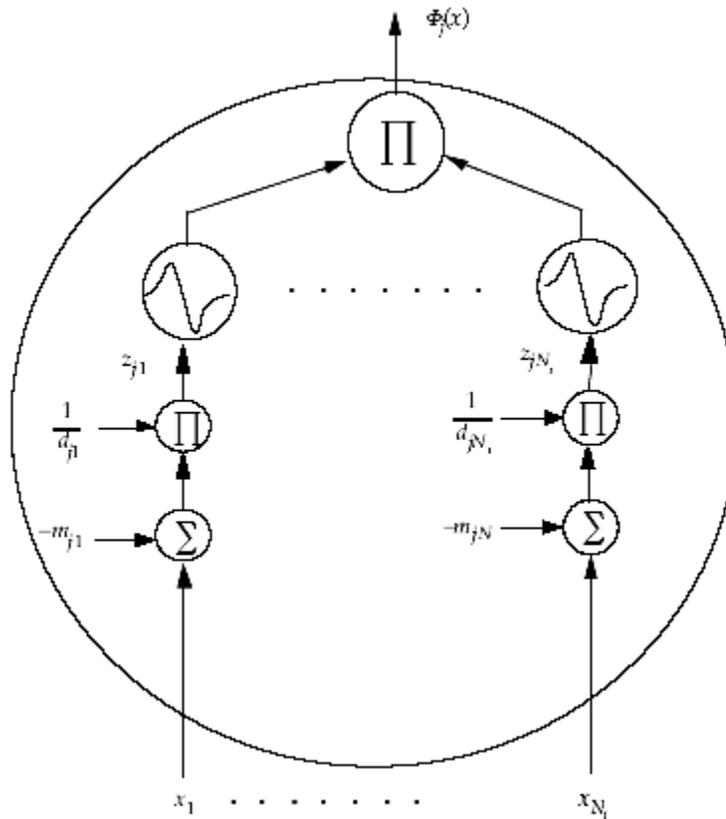


Figure 4.2: Représentation graphique d'une ondelette multidimensionnelle séparable. m_{ij} : translation, d_{ij} : dilatation [Ouss98].

Le réseau d'ondelettes peut être considéré comme constitué de trois couches. Une première couche avec N_i entrée(s), une couche cachée constituée par N_w ondelettes et un sommateur (ou neurone linéaire) de sortie recevant les sorties pondérées des ondelettes multidimensionnelles et la partie affine. Ce réseau est illustré par la figure 4.3.

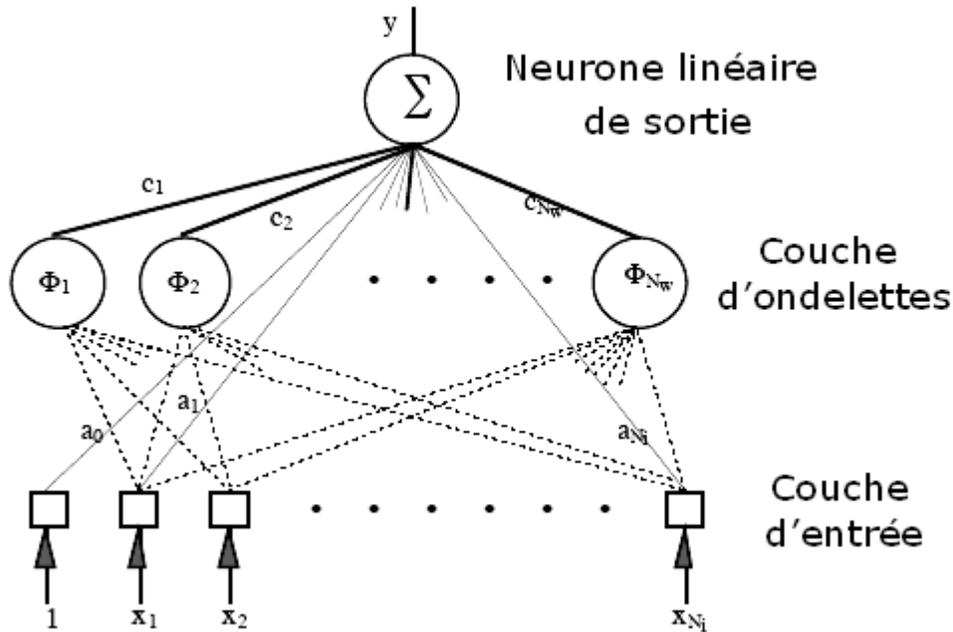


Figure 4.3: Représentation graphique d'un réseau d'ondelettes [zhan92].

Les réseaux d'ondelettes sont des réseaux de fonctions non linéaires paramétrées, où le vecteur Θ_i est constitué par les translations et les dilatations de l'ondelette multidimensionnelle. Plusieurs choix d'ondelettes sont possibles. En effet, plusieurs familles d'ondelettes existent. Les ondelettes les plus connues (et aussi les plus anciennes) sont certainement celles qui constituent le système de Haar que l'on présentera dans la suite du chapitre, dans le contexte d'ondelettes orthogonales.

Les fonctions du système de Haar n'étant pas dérivables, il n'est pas possible d'appliquer aux réseaux de telles ondelettes. Les ondelettes que nous allons utiliser pour la construction de réseaux sont celles issues des travaux de I. Daubechies. On parle dans ce cas d'ondelettes de la famille de Daubechies.

Une des ondelettes-mères que nous utilisons dans ce mémoire est la dérivée première de la fonction gaussienne [zhan92]. C'est l'une des ondelettes les plus utilisées [torr95]. Elle est définie par :

$$\phi(x) = \pm x e^{-\frac{1}{2}x^2} \quad \text{et} \quad \phi_j(x) = \pm \frac{1}{\sqrt{d_j}} \left(\frac{x \pm m_j}{d_j} \right) \exp \left(\pm \frac{1}{2} \left(\frac{x \pm m_j}{d_j} \right)^2 \right) \quad 4.5$$

Le graphe de cette fonction est représenté sur la figure 4.4.

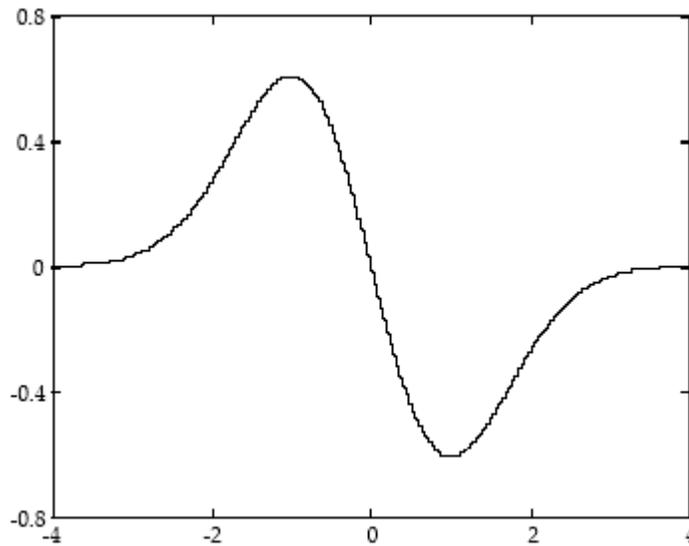


Figure 4.4: Graphe d'une ondelette.

Cette ondelette peut être considérée comme une forme dérivable des ondelettes du système de Haar, de même que la tangente hyperbolique (utilisée comme fonction d'activation des réseaux de neurones) est une forme dérivable de la fonction signe.

Une autre ondelette-mère que l'on rencontre souvent dans la bibliographie est la dérivée seconde de la fonction gaussienne [cann95], [baro97]. Son expression est :

$$\phi(x) = (x^2 \pm 1)e^{-\frac{1}{2}x^2} \quad 4.6$$

Elle est appelée "ondelette chapeau mexicain". Son graphe est le suivant :

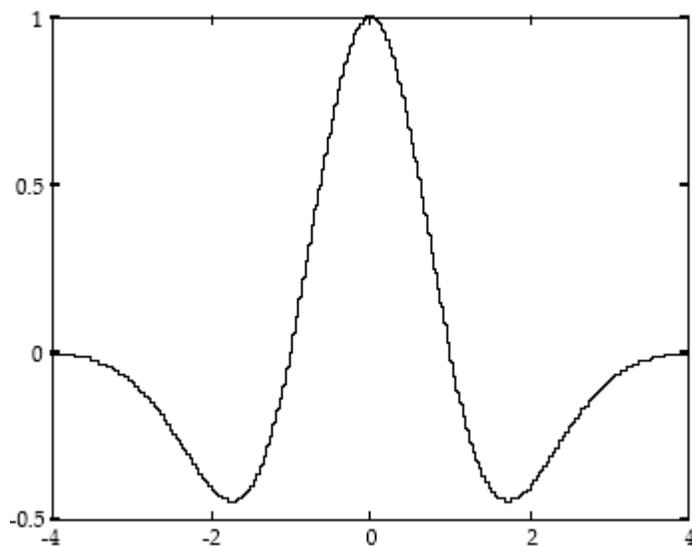


Figure 4.5: Graphe de l'ondelette "chapeau mexicain".

4.3.3 Comparaison entre Réseaux d'ondelettes et réseaux de neurones

La principale ressemblance entre les réseaux de neurones à fonctions dorsales, et les réseaux d'ondelettes, réside dans le fait que les deux réseaux calculent une combinaison linéaire, à paramètres ajustables, de fonctions non linéaires dont la forme dépend de paramètres ajustables (translations et dilatations).

Les différences essentielles entre ces deux types de réseaux sont les suivantes :

- Contrairement aux fonctions dorsales, les ondelettes sont des fonctions qui décroissent rapidement, et tendent vers zéro dans toutes les directions de l'espace. Elles sont donc locales si dj est petit ;
- contrairement aux fonctions dorsales, la forme de chaque ondelette monodimensionnelle est déterminée par deux paramètres ajustables (translation et dilatation) qui sont des paramètres structurels de l'ondelette.
- chaque ondelette monodimensionnelle possède deux paramètres structurels, donc, pour chaque ondelette multidimensionnelle, le nombre de paramètres ajustables est le double du nombre de variables.

Pour comparer la complexité des réseaux, deux éléments sont importants :

- Le nombre de paramètres ajustables
- le nombre d'opérations élémentaires à effectuer.

Nous utiliserons les notations suivantes :

- ❖ Nombre d'entrées : N_i (entrée constante non comprise), déciles extraits de la DSP des contractions.
- ❖ Nombre de fonctions : N_w (pour les ondelettes), N_c (pour les fonctions dorsales).
- ❖ Nombre de sorties : Une.

Le tableau 4.1 présente les éléments de comparaison pour ces deux méthodes:

	Réseaux de fonctions dorsales	Réseaux de fonctions ondelettes
Nombre de fonctions.	N_c	N_w
Nombre de paramètres.	$N_i + 2 N_c + 1 \pm 1$	$2 N_i N_w + N_w + N_i + 1$
Nombre d'opérations pour le calcul de la sortie.	$N_c 2 N_i + 3 + N_i + 1$	$3 N_w N_i + 2 + N_i + 1$

Tableau 4.1. Eléments de comparaison entre réseaux d'ondelettes et de fonctions dorsales.

On entend par "opération" les opérations mathématiques élémentaires, c'est-à-dire addition, multiplication ou division. Étant données les propriétés de la fonction exponentielle, et pour chacun des deux types de réseaux, il y a autant de fonctions exponentielles à calculer que de neurones cachés ou d'ondelettes multidimensionnelles dans le réseau.

Afin de savoir lequel des deux types de réseaux est le plus économique en termes de nombre d'opérations nécessaires pour le calcul de la sortie, il est possible de calculer la différence entre les deux résultats de la dernière ligne du tableau ci-dessus. En effet, à nombre de fonctions égales ($N_w = N_c$), la différence entre les nombre d'opérations pour les deux types de réseaux est égale à $N_w (N_i + 3)$.

Le nombre d'opérations effectuées lors du calcul de la sortie avec un réseau d'ondelettes est donc toujours supérieur à celui effectué par un réseau de fonctions dorsales ayant le même nombre d'entrées et de fonctions.

4.4 Apprentissage des réseaux d'ondelettes non bouclés

4.4.1 Calcul du gradient de la fonction de coût

Les coefficients du réseau peuvent être divisés en deux classes :

- les paramètres structurels des fonctions, c'est-à-dire les translations et les dilatations.
- les coefficients de pondérations c_j et les coefficients a_k de la partie affine.

Deux possibilités s'offrent à nous pour la construction du réseau :

- choisir les paramètres structurels dans un ensemble de valeurs discrètes
- considérer ces paramètres comme ceux d'un réseau de neurones classique et utiliser une technique d'optimisation pour en faire une estimation.

Discrétiser le domaine des translations et des dilatations signifie qu'on effectue la construction de réseaux d'ondelettes suivant une approche fondée sur la transformée en ondelettes discrète. Cette question sera étudiée en détail dans la suite de ce chapitre.

Dans ce qui suit, nous allons adopter la seconde possibilité, et faire appel aux techniques d'optimisation non linéaire.

Rappelons que l'apprentissage consiste à minimiser la fonction de coût suivante :

$$J(\theta) = \frac{1}{2} \sum_{n=1}^N (y_p^n \pm y^n)^2 = \frac{1}{2} \sum_{n=1}^N (e^n)^2 \quad 4.7$$

avec

$$y^n = \psi(x^n, \theta) = \sum_{j=1}^{N_w} c_j \Phi_j(x^n, m_j, d_j) + \sum_{k=0}^{N_i} a_k x_k^n \quad \text{avec } x_0^n = 1 \quad 4.8$$

où y_p^n est la sortie désirée correspondant à l'exemple n , y^n est la sortie du réseau d'ondelettes pour l'exemple n , et

$$X^n = \{X_1^n, \dots, X_{N_i}^n\} \quad 4.9$$

est le vecteur des entrées.

θ est le vecteur regroupant l'ensemble des paramètres ajustables :

$$\theta = \{m_{jk}, d_{jk}, c_j, a_k, a_0\} \quad j = 1, \dots, N_w \quad \text{et} \quad k = 1, \dots, N_i \quad 4.10$$

Les techniques d'optimisation utilisées nécessitent le calcul du vecteur gradient de la fonction de coût par rapport au vecteur des paramètres ajustables. Son expression est :

$$\frac{\partial J}{\partial \theta} = \pm \sum_{n=1}^N e^n \frac{\partial y^n}{\partial \theta} \quad 4.11$$

où $\partial y_n / \partial \theta$ est la valeur du gradient de la sortie du réseau par rapport aux paramètres θ au point $x=x^n$:

$$\frac{\partial y^n}{\partial \theta} = \frac{\partial y}{\partial \theta} \Big|_{x=x^n} \quad 4.12$$

Calculons à présent la dérivée de la sortie par rapport à chacun des paramètres du réseau.

• Pour les coefficients directs $\{ak\}$:

$$\frac{\partial y^n}{\partial a_k} = x_k^n \quad k = 1, \dots, N_i \quad 4.13$$

• Pour les pondérations des ondelettes $\{cj\}$:

$$\frac{\partial y^n}{\partial c_j} = \Phi_j(x^n) \quad k = 1, \dots, N_i \quad \text{et} \quad j = 1, \dots, N_w \quad 4.14$$

• Pour les translations $\{mjk\}$:

$$\frac{\partial y^n}{\partial m_{jk}} = \pm \frac{c_j}{d_{jk}} \frac{\partial \Phi_j}{\partial z_{jk}} \Big|_{x=x^n} \quad k = 1, \dots, N_i \quad \text{et} \quad j = 1, \dots, N_w \quad 4.15$$

• Pour les dilatations $\{djk\}$:

$$\frac{\partial y^n}{\partial d_{jk}} = \pm \frac{c_j}{d_{jk}^2} z_{jk}^n \frac{\partial \Phi_j}{\partial z_{jk}} \Big|_{x=x^n} \quad k = 1, \dots, N_i \quad \text{et} \quad j = 1, \dots, N_w \quad 4.16$$

$\frac{\partial \phi_j}{\partial z_{jk}} \Big|_{x=x^n}$ est la valeur de la dérivée partielle de l'ondelette multidimensionnelle par rapport à la variable z_{jk} au point $x=x^n$. Étant donné la relation (4.3), cette dérivée partielle vaut :

$$\frac{\partial \phi_j}{\partial z_{jk}} \Big|_{x=x^n} = \phi(z_{j1}^n) \phi(z_{j2}^n) \dots \phi'(z_{jk}^n) \dots \phi(z_{jN_i}^n) \quad 4.17$$

avec $\phi'_{z_{jk}}$ la dérivée au point $x=x^n$ de l'ondelette scalaire, c'est-à-dire :

$$\phi'_{z_{jk}} = \frac{d\phi(z)}{dz} \Big|_{z=z_{jk}^n} \quad 4.18$$

4.4.2 Initialisation des paramètres du réseau

Une fonction d'ondelettes monodimensionnelle est définie sur tout l'ensemble \mathbf{R} , mais l'essentiel de sa contribution s'étend sur un intervalle centré autour de la valeur de la translation et dont la longueur dépend du paramètre de dilatation.

Dans le cas de réseaux de neurones à fonctions dorsales, l'initialisation des paramètres du réseau est généralement effectuée de manière aléatoire, de telle manière que le potentiel de

chaque neurone caché soit suffisamment petit pour que les sorties des neurones se trouvent dans la partie linéaire de la sigmoïde.

Les ondelettes étant des fonctions à décroissance rapide, une initialisation aléatoire des paramètres de translation et de dilatation serait très inefficace : en effet, si les translations sont initialisées à l'extérieur du domaine contenant les exemples, ou si les dilatations choisies sont trop petites, la sortie de l'ondelette est pratiquement nulle, de même que sa dérivée.

L'algorithme d'adaptation des paramètres étant fondé sur une technique de gradient, il est dans ce cas inopérant. Une attention particulière doit donc être portée à cette phase d'initialisation des paramètres. Nous proposons ici une procédure d'initialisation simple, qui prend en considération le domaine où sont répartis les exemples de l'ensemble d'apprentissage.

Soit $[\alpha_k, \beta_k]$ l'intervalle contenant les k ème composantes des vecteurs d'entrée des exemples.

On initialise les translations m_{jk} ($j = 1, \dots, N_w$) au centre de l'intervalle $[\alpha_k, \beta_k]$:

$$m_{jk} = \frac{\alpha_k + \beta_k}{2} \quad \text{avec } j=1, \dots, N_w \quad 4.19$$

Les paramètres de dilatation sont choisis de telle manière que les variations de l'ondelette s'étendent sur tout l'intervalle $[\alpha_k, \beta_k]$. Cette condition est remplie avec le choix suivant :

$$d_{jk} = 0,2(\alpha_k \pm \beta_k) \quad \text{avec } j=1, \dots, N_w \quad 4.20$$

Cette procédure est valable notamment pour l'ondelette mère illustrée par la figure 4.4 que nous allons utiliser dans nos exemples.

Reste la question de l'initialisation des coefficients de pondération des ondelettes (c_j avec $j = 1, \dots, N_w$) et de ceux de la partie affine ak avec $k = 1, \dots, N_i$.

L'initialisation de ces coefficients est moins importante, pour le déroulement de l'apprentissage, que celle des paramètres structurels ; ils sont initialisés de manière aléatoire, uniformément répartis dans l'intervalle $[-10^{-2}; +10^2]$.

Cette procédure ne nécessite pratiquement pas de calcul ; elle est très simple à mettre en œuvre. Le fait que, pour $j = 1, \dots, N_w$ toutes les translations soient initialisées à la même valeur (ainsi que les dilatations) peut laisser penser qu'elles vont évoluer de manière identique si l'on effectue plusieurs apprentissages successifs. Une telle situation est évitée par le fait que les pondérations de chacune des ondelettes du réseau sont initialisées différemment de par leur tirage aléatoire.

Néanmoins, cette procédure d'initialisation présente un inconvénient : elle utilise peu les propriétés des ondelettes. En effet, on peut imaginer qu'on puisse mettre au point une technique d'initialisation qui utilise plus l'information apportée par les paramètres structurels, afin que la fonction de coût soit au voisinage d'un minimum avant d'effectuer l'apprentissage proprement dit. Une telle procédure est proposée dans la suite ; elle est utilisable pour des réseaux d'ondelettes issus de la transformée en ondelettes discrète.

4.5. Réseaux issus de la transformée en ondelettes discrètes

4.5.1 Introduction

Dans le contexte précédent, nous avons présenté des procédures d'apprentissage pour des réseaux d'ondelettes fondés sur la transformée en ondelettes continue. L'application de ces procédures est possible, car les paramètres des fonctions (et en particulier ceux des fonctions ondelettes) sont choisis de manière continue dans l'ensemble des réels. Dans la partie présente, nous proposons des procédures de construction de réseaux d'ondelettes dont les paramètres sont à valeurs discrètes. On utilise donc ici la transformée en ondelettes discrète.

Nous présenterons tout d'abord le principe de la transformée en ondelettes discrète. Nous examinerons ensuite les méthodes que l'on peut mettre en œuvre pour la construction de réseaux d'ondelettes à paramètres discrets pour les applications de classification. Nous verrons que, comme on peut le prévoir, cette restriction donne moins de souplesse dans le choix d'un modèle que l'approche fondée sur la transformée continue. De plus, pour des raisons inhérentes aux techniques utilisées, certains modèles-hypothèses possibles avec la transformée en ondelettes continue ne sont pas envisageables dans un contexte de transformée discrète.

Comme pour les ondelettes à paramètres continus, nous proposerons une procédure d'initialisation utilisant la transformée discrète. L'apprentissage sera ensuite effectué suivant les algorithmes présentés dans la partie précédente. Ainsi, les deux approches jouent un rôle complémentaire pour la construction d'un réseau d'ondelettes.

Une transformée en ondelettes est dite discrète lorsque les valeurs des translations et des dilatations sont à valeurs discrètes (pas nécessairement entières). Soit ϕ une ondelette mère (qui peut être la même que celle utilisée pour la transformée continue). Une famille Ω de fonctions obtenue à partir de ϕ peut être exprimée de la manière suivante :

$$\Omega(\alpha, \beta, x) = \left\{ \alpha^{m/2} \phi(\alpha^m x \pm n\beta), (m, n) \in \mathbf{Z}^2 \right\} \quad 4.21$$

où \mathbf{Z} est l'ensemble des entiers relatifs. Notons que α et β sont des paramètres réels fixes qui définissent, avec ϕ , la famille Ω : α détermine l'échelle des dilatations et β détermine le pas des translations. Une famille d'ondelettes est donc entièrement connue par la donnée du triplet (ϕ, α, β) . Un membre de cette famille (c'est-à-dire une fonction) est désignée par le couple (m, n) . Pour cette raison, on désignera dorénavant une ondelette de la famille de Ω ayant comme paramètres (m, n) par: $\phi_{m,n}(x)$.

Suivant la relation (4.21), nous avons :

$$\phi_{m,n}(x) = \alpha^{m/2} \phi(\alpha^m x \pm n\beta) \quad 4.22$$

4.5.2 Structures obliques et bases d'ondelettes orthonormales

Ondelettes à variables continues

Il a été démontré dans [daub92] qu'une famille d'ondelettes $\phi_{m,n}(x)$ comme celles décrites plus haut possède la propriété de structure oblique de l'ensemble $L2(\mathbf{R})$. Rappelons cette propriété : étant donné une fonction f de carré sommable, il existe deux constantes c et C positives et de valeurs finies telles que l'inégalité suivante soit vérifiée :

$$c|f|^2 \leq \sum_{\phi_{m,n} \in \Omega} |\langle \phi_{m,n}, f \rangle|^2 \leq C|f|^2 \quad 4.23$$

Les valeurs de ces deux constantes (dites "limites de la structure oblique") donnent une indication sur la qualité de l'approximation de la fonction f par la famille d'ondelettes Ω [zhan92] : en particulier, pour une somme finie d'éléments de Ω , plus ces constantes sont proches de 1, meilleure est la qualité de l'approximation. A partir des structures obliques, on définit les *structures obliques étroites* ("tight frames" en anglais) de la manière suivante [pati93] :

(a) Une structure oblique ayant des limites égales (c'est-à-dire $c = C$) est dite une *structure oblique étroite*.

(b) Une *structure oblique étroite* dont les éléments sont normés et ayant $c = C = 1$ est une base orthonormale.

Nous allons maintenant présenter un exemple de famille d'ondelettes qui constitue une base de fonctions orthogonales.

Avec $\alpha = 2$, $\beta = 1$ et un choix adéquat de l'ondelette mère (c'est-à-dire de ϕ) il est possible de construire une base d'ondelettes orthonormales. Elle peut s'exprimer de la façon suivante :

$$\Omega(2,1) = \left\{ 2^{m/2} \phi(2^m x \pm n), (m,n) \in \mathbf{Z}^2 \right\} \quad 4.24$$

Les translations et les dilatations peuvent être obtenues en utilisant la réécriture de la relation (4.23). La famille d'ondelettes orthonormales la plus connue est une base appelée "Système de Haar". L'ondelette mère est une fonction définie par morceaux :

$$\phi(x) = \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2}[\\ \pm 1 & \text{si } x \in [0, \frac{1}{2}] \\ 0 & \text{ailleurs} \end{cases} \quad 4.25$$

La figure 4.6 illustre trois ondelettes de cette famille :

$$\phi_{0,0}(x) = \phi(x) \quad (\text{Figure 4.6.a}) \quad 4.26$$

$$\phi_{\pm 1,0}(x) = \frac{\sqrt{2}}{2} \phi\left(\frac{x}{2}\right) \quad (\text{Figure 4.6.b}) \quad 4.27$$

$$\phi_{\pm 1,1}(x) = \frac{\sqrt{2}}{2} \phi\left(\frac{x \pm 2}{2}\right) \quad (\text{Figure 4.6.c}) \quad 4.28$$

Notons que le centre de la fonction (c'est-à-dire le paramètre de translation) n'est pas le centre de symétrie du graphe de la fonction mais la limite gauche de la partie non nulle de la fonction. Par exemple pour l'ondelette mère, le paramètre de translation est égal à 0 et non pas à $\frac{1}{2}$.

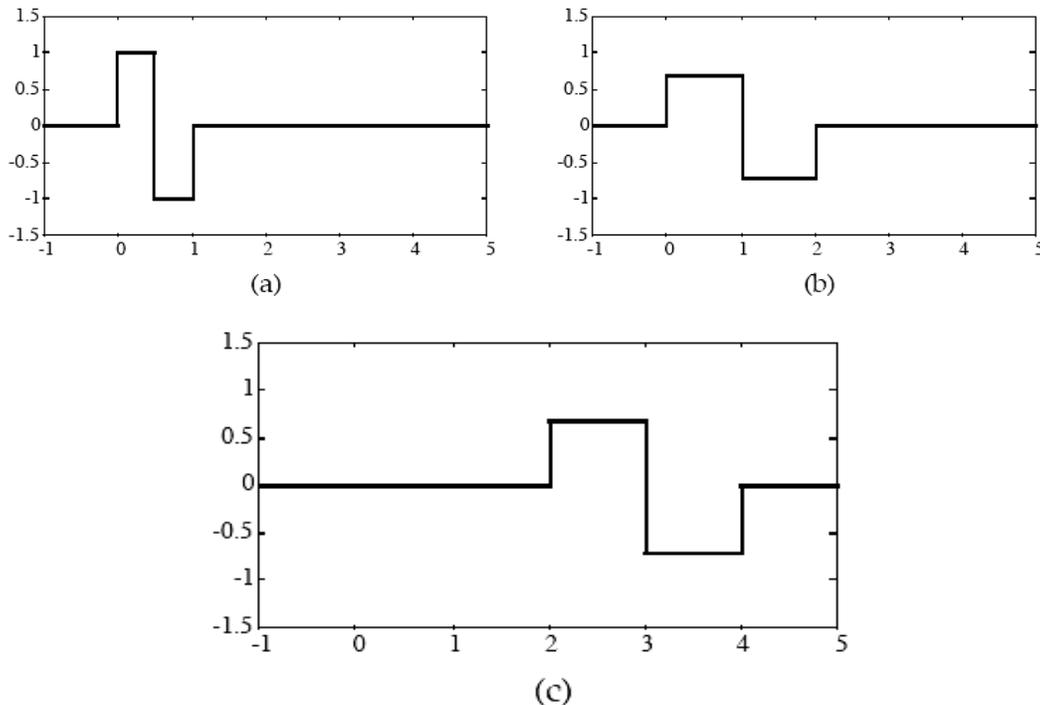


Figure 4.6. Trois ondelettes du Système de Haar.

On peut facilement vérifier que ces fonctions possèdent bien la propriété d'orthogonalité. La norme d'une fonction étant son carré scalaire, elle est égale ici à $(\phi_{m,n}, \phi_{m,n}) = \alpha^m \alpha^{\pm m} = 1$. Le système de Haar est donc une famille d'ondelettes orthonormales. Le système de Haar est l'une des rares familles orthonormales (dont l'ondelette mère possède une expression simple) que l'on peut trouver dans la littérature. Cette rareté est principalement due à la difficulté du choix de l'ondelette mère.

Ondelettes à variables discrètes

Dans tout ce qui précède, nous avons considéré l'orthogonalité de fonctions de variables continues, avec le produit scalaire défini par la relation $(\phi_{m,n}, \phi_{m,n}) = \alpha^m \alpha^{\pm m} = 1$. Dans le cas où l'une des fonctions du produit scalaire est connue en un nombre fini de points (typiquement pour un problème d'approximation de fonction) la relation devient une sommation discrète :

$$\langle \phi_{m,n}, \phi_{m',n'} \rangle = \sum_{\substack{i=1 \\ x^i \in A}}^N \phi_{m,n}(x^i) \phi_{m',n'}(x^i) \quad 4.29$$

où A désigne l'ensemble des points et N leur nombre. Le produit scalaire de deux fonctions, donc leur éventuelle orthogonalité, dépend ainsi du choix des points de A . Il est facile de vérifier (par exemple sur le système de Haar) qu'un choix quelconque du nombre de points et de leur distribution ne conduit pas à la conservation de l'orthogonalité.

En revanche, lorsqu'on veut modéliser un processus (c'est-à-dire approcher une fonction de régression hypothétique et *inconnue*, ou une application de classification), le choix des points d'échantillonnage est rarement libre : le concepteur du modèle ne dispose souvent que d'une base de données existante, ou bien le choix des mesures peut être contraint par les conditions normales de fonctionnement du processus, que le modélisateur n'a pas le droit de modifier librement (par exemple pour la classification de données d'un processus industriel opérationnel).

Choix de l'ondelette mère

Comme nous l'avons souligné plus haut dans ce paragraphe, la propriété d'orthogonalité est obtenue par un bon choix des paramètres α et β et mais aussi de l'ondelette mère ϕ . En effet, toute ondelette mère ne permet pas la construction de bases orthonormales. Le système de Haar constitue une base de fonctions orthogonales, mais ces fonctions ne sont pas régulières, ce qui rend leur utilisation malaisée en tant que classificateur.

En effet, que ce soit en approximation de fonction ou en modélisation de processus ou en classification, on utilise des familles de fonctions régulières. Un des objectifs de la théorie des bases d'ondelettes est la construction de systèmes ayant les mêmes propriétés que le système de Haar, mais dont l'ondelette mère serait régulière. Ceci permettrait ainsi l'utilisation de ces bases pour résoudre des problèmes de classification.

Pour construire de manière systématique des ondelettes orthogonales, il convient de généraliser le principe de l'approximation multirésolution à celui de l'analyse multirésolution. Plusieurs mathématiciens se sont penchés sur la construction d'ondelettes orthogonales s'appuyant sur la théorie de la multirésolution. Les premières ont été proposées par Meyer. Malheureusement, ces ondelettes mères orthogonales ne possèdent pas d'expression analytique simple, ce qui rend difficile l'utilisation de ces fonctions pour des problèmes de classification. On se contentera donc, dans la suite de cette étude, de faire appel uniquement à des familles d'ondelettes constituant des structures obliques et particulièrement des structures obliques étroites, plutôt qu'à des bases orthonormales. Les structures obliques peuvent être considérées comme des bases d'ondelettes redondantes.

Leur principal inconvénient, par rapport à des bases orthonormales proprement dites, réside dans la perte de l'unicité de la représentation d'une fonction et la signification que peut avoir la pondération dans le cas d'ondelettes orthogonales.

Dans la suite de ce chapitre, nous allons nous intéresser aux familles d'ondelettes issues de la transformée discrète en tant qu'outil pour l'application de classification, en définissant des réseaux d'ondelettes fondés sur la transformée discrète.

4.5.3 Réseaux fondés sur la transformée discrète

Lorsque la transformation en ondelettes discrète, est effectuée de manière appropriée (choix adéquat de la discrétisation des paramètres de translation et de dilatation comme décrit ci-dessus), alors la transformation inverse existe. De ce fait, toute fonction f de $L2(\mathbf{R})$ peut être représentée comme une somme d'éléments d'une base orthonormale ou d'une structure oblique d'ondelettes (comme celles décrites plus haut dans ce chapitre) suivant la relation:

$$f(x) = \sum_{(m,n) \in \mathbb{Z}^2} c_{mn} \phi_{m,n}(x) \quad 4.30$$

Nous définissons un réseau d'ondelettes fondé sur la transformée en ondelettes discrète comme une somme (relation 4.31), à laquelle on ajoute des termes directs. La sortie d'un tel réseau est donc donnée par la relation suivante :

$$y = \psi(x) = \sum_{j=1}^{N_w} c_j \Phi_j(x) + \sum_{k=0}^{N_i} a_k x_k \quad 4.31$$

où N_w est le nombre d'ondelettes et N_i le nombre d'entrées. $\phi_j(x)$ est une ondelette multidimensionnelle obtenue par produit de toutes les ondelettes suivant chacune des entrées. Ces réseaux sont, du point de vue de leur structure, identiques à ceux définis avec la transformée continue. La principale différence réside dans les méthodes de détermination des translations et des dilatations.

4.6 Techniques de construction de réseaux d'ondelettes discrettes

4.6.1 Impossibilité d'utiliser les techniques de gradient

L'architecture des réseaux d'ondelettes fondés sur la transformée discrète ayant été définie, nous nous posons la question de la construction d'un modèle, constitué d'un réseau d'ondelettes, d'un processus donné. Comme précédemment, nous considérons que les seules connaissances sur le processus sont constituées d'une (ou plusieurs) séquence(s) d'entrées et de sorties mesurées (si le processus est réel).

Les paramètres à déterminer pour la construction du réseau sont:

- le nombre d'ondelettes nécessaires pour atteindre une performance voulue.

- les valeurs à donner aux différents paramètres du réseau : paramètres structurels, pondérations des ondelettes et termes directs.

La difficulté essentielle réside dans la détermination des paramètres du réseau. Ceux-ci prenant des valeurs discrètes, la minimisation d'un coût utilisant le gradient n'est pas envisageable. En revanche, on peut tirer profit du fait que les paramètres prennent des valeurs discrètes pour concevoir des méthodes de sélection des ondelettes dans un ensemble (bibliothèque) d'ondelettes discrètes. La performance du modèle ainsi conçu dépend du choix initial des ondelettes de la bibliothèque, et d'une sélection judicieuse dans cette bibliothèque.

4.6.2 Approches utilisant une procédure de sélection

Technique fondée sur la construction de structures obliques étroites

Étant donné les limites théoriques auxquelles on se heurte pour la construction de réseaux d'ondelettes orthogonales (il n'existe pas d'expression analytique simple pour des ondelettes mères engendrant des bases orthonormales), on se propose ici d'utiliser des structures obliques. La question qui se pose alors est le choix des paramètres α et β . On construit la bibliothèque à l'aide d'une *structure oblique étroite*. Les paramètres α et β sont alors respectivement égaux à 2 et à 1 [judi94], [zhan97]. La bibliothèque est construite avec quatre ou cinq dilatations différentes. L'ondelette la plus large est celle dont le support a la taille du domaine des exemples. Les ondelettes retenues sont celles dont les centres sont à l'intérieur de ce domaine. Une méthode de sélection constructive ou destructive est ensuite appliquée aux ondelettes retenues dans la bibliothèque pour déterminer celles qui sont les plus significatives pour modéliser le processus étudié. Zhang en 1993 a proposé d'appliquer une première réduction de la bibliothèque en éliminant les ondelettes comportant peu ou pas d'exemples sur leurs supports [zhan93] et [zhan97]. Ces situations sont particulièrement fréquentes pour des modèles à plusieurs entrées où les exemples ne sont pas répartis de manière uniforme. Cette technique présente l'avantage de procéder à une construction de la bibliothèque de manière simple, qui nécessite peu de calculs.

4.7 Proposition d'une procédure de construction de réseaux et d'initialisation de l'apprentissage

Dans ce paragraphe, nous présentons une méthode de sélection d'ondelettes :

1. pour la construction de réseaux d'ondelettes fondés sur la transformée discrète.
2. pour l'initialisation, avant apprentissage, des translations et des dilatations de réseaux d'ondelettes fondés sur la transformée continue.

L'étape de la construction de la bibliothèque est fondée sur la théorie des *structures obliques étroites* d'ondelettes. Elle est donc semblable à celle présentée dans [zhan97] excepté par le fait que, contrairement aux réseaux présentés dans cette référence, chaque ondelette a des dilatations différentes suivant différentes entrées. Ce choix présente l'avantage d'enrichir la bibliothèque, et d'obtenir une meilleure performance pour un nombre de fonctions donné.

L'inconvénient introduit par ce choix concerne la taille de la bibliothèque. Une bibliothèque d'ondelettes ayant des dilatations différentes pour chaque entrée est plus volumineuse que

celle dont les ondelettes possèdent la même dilatation suivant toutes les entrées. Ceci implique un coût de calcul plus élevé pendant l'étape de sélection. Néanmoins, la sélection d'ondelettes est souvent plus courte que l'apprentissage des dilatations et translations par les techniques de gradient utilisées pour les ondelettes à paramètres continus ; le coût supplémentaire introduit par des dilatations différentes peut donc être acceptable.

Dans ce qui suit, nous exposerons successivement la construction de la bibliothèque, la sélection des ondelettes dans cette bibliothèque, puis les procédures de construction et d'initialisation. On insistera particulièrement sur l'étape de construction de la bibliothèque. En effet, dans la littérature, la description pratique de cette étape est très rarement abordée en détail.

4.7.1 Description de la procédure de construction de la bibliothèque

Rappelons qu'on se propose de construire une bibliothèque d'ondelettes candidates pour notre application de classification. On dispose d'une séquence d'exemples répartis dans l'intervalle $[a, b]$. On considérera d'abord un modèle à une seule entrée. La généralisation pour un modèle à N_i entrées sera décrite par la suite.

Famille engendrant la bibliothèque pour un modèle à une entrée

Rappelons que la bibliothèque est engendrée à partir de la famille suivante :

$$\mathcal{W} = \left\{ \phi_{m,n}^{(\alpha, \beta)}(x) \mid m \in \mathbb{Z} \right\} \quad 4.32$$

Étant donné que la construction est fondée sur une *structure oblique étroite*, nous avons $\alpha = 2$ et $\beta = 1$.

Une ondelette de cette famille avec les paramètres m et n s'exprime en fonction de l'ondelette mère de la façon suivante :

$$\phi_{m,n}(x) = 2^{m/2} \phi \left(\frac{x \pm 2_n^{-m}}{2^{-m}} \right) \quad 4.33$$

C'est donc une ondelette ayant pour centre 2_n^{-m} et pour dilatation 2^{-m} .

Choix des dilatations

Plus m est grand, plus le nombre d'ondelettes nécessaires est grand. Pour des raisons de taille de la bibliothèque, on se limite à trois dilatations successives (c'est-à-dire trois valeurs entières successives du paramètre m). Il suffit donc de choisir la plus grande dilatation, ou la plus petite, pour que les deux autres soient déterminées. Pour l'ondelette mère que nous utilisons dans nos exemples $\phi(x) = \pm x e^{\pm \frac{x^2}{2}}$, la valeur de la dilatation assurant que

l'ondelette, centrée au milieu de l'intervalle $\left(\text{en } \frac{a+b}{2} \right)$, a sa partie utile aussi large que le domaine $[a, b]$ est : $0,2(b-a)$. Cette valeur de la dilatation est obtenue en estimant que la partie utile de l'ondelette coïncide avec le domaine où la sortie de l'ondelette est supérieure à 0.1 ou inférieure à -0.1. Elle est solution d'une équation non algébrique donnée par :

$$\frac{b \pm a}{2d} \exp\left(\pm \frac{1}{2} \left(\frac{b \pm a}{2d}\right)^2\right) = 0.1 \quad 4.34$$

Notons que cette propriété a été utilisée pour l'initialisation des réseaux fondés sur la transformée continue. On considère que cette valeur est celle de la plus grande dilatation. Les dilatations suivantes sont donc plus petites. Ceci peut se traduire par la relation suivante :

$$2^{-m} \leq 0.2(b - a) \quad 4.35$$

ce qui est équivalent à :

$$m \geq \pm \frac{\text{Log}(0.2(b \pm a))}{\text{Log}2} \quad 4.36$$

m est entier, alors que le second membre de cette inégalité ne l'est probablement pas. En pratique, la plus petite valeur de m à considérer (elle correspond à la plus grande dilatation) sera :

$$\left[\pm \frac{\text{Log}(0.2(b \pm a))}{\text{Log}2} \right] \quad 4.37$$

où l'opérateur $[]$ désigne la fonction partie entière.

Les trois valeurs du paramètre m que l'on utilise pour la construction de la bibliothèque sont donc :

$$\left\{ \left[\pm \frac{\text{Log}(0.2(b \pm a))}{\text{Log}2} \right] + 1, \left[\pm \frac{\text{Log}(0.2(b \pm a))}{\text{Log}2} \right] + 2, \left[\pm \frac{\text{Log}(0.2(b \pm a))}{\text{Log}2} \right] + 3 \right\} \quad 4.38$$

Choix des translations

Pour une dilatation donnée, on retient dans la bibliothèque toutes les ondelettes dont les centres sont à l'intérieur du domaine $[a, b]$. Pour une valeur de m donnée, cette condition peut s'exprimer de la manière suivante :

$$a \leq 2^{-m} n \leq b \quad 4.39$$

Étant donné que l'on cherche ici à déterminer les valeurs possibles pour n , la condition précédente est équivalente à :

$$2^m a \leq n \leq 2^m b \quad 4.40$$

Là aussi, étant donné que n est un nombre entier, les valeurs possibles sont (avec $[]$ l'opérateur partie entière) :

$$\{[2^m a] + 1, [2^m a] + 2, \dots, [2^m b]\} \quad 4.41$$

En pratique, chaque fois que m est augmenté d'une unité, le nombre d'ondelettes ajoutées par $m+1$ à la bibliothèque est double de celui apporté par m . La bibliothèque est donc construite suivant un schéma pyramidal.

Cas des bibliothèques pour modèles à plusieurs entrées

Dans le cas d'un problème multidimensionnel, ce calcul est effectué pour chacune des entrées. Étant donné qu'une ondelette multidimensionnelle est le produit des ondelettes scalaires, le cardinal de la bibliothèque est égal au produit du nombre d'ondelettes suivant chacune des entrées.

4.7.2 Méthode de sélection par orthogonalisation

La bibliothèque étant construite, une méthode de sélection est ensuite appliquée afin de déterminer les ondelettes les plus significatives pour modéliser le processus considéré. Soit M_w le nombre d'éléments dans la bibliothèque et N_w le nombre d'ondelettes dans le réseau. Pour sélectionner les N_w ondelettes qui permettent de constituer le modèle dont l'EQMA (l'erreur quadratique moyenne d'apprentissage) est la plus faible possible, l'idéal serait de calculer les EQMA obtenues avec tous les sous-ensembles de cardinal N_w qu'on peut former à partir d'un ensemble de cardinal M_w . Ce nombre de sous-ensembles est généralement très grand. De ce fait, on a recours à une méthode de sélection qui présente un moindre coût du point de vue du volume de calculs nécessaires.

La technique de sélection qu'on utilise effectue un classement des ondelettes de la bibliothèque sur la base de la procédure d'orthogonalisation de Gram-Schmidt. Cette procédure est proposée dans plusieurs références [Chen89], [Zhan93] dans le cadre de la construction de réseaux d'ondelettes.

Soit une séquence d'apprentissage formée de N exemples. On considère une bibliothèque contenant M_w ondelettes candidates. A chaque ondelette ϕ_j on associe un vecteur dont les composantes sont les valeurs de cette fonction suivant les exemples de la séquence d'apprentissage. On constitue ainsi une matrice P dont l'expression est :

$$P = \begin{pmatrix} \Phi_1(x_1) & \Phi_2(x_1) & \dots & \dots & \dots & \Phi_{M_w}(x_1) \\ \Phi_1(x_2) & \Phi_2(x_2) & & & & \vdots \\ \Phi_1(x_3) & \vdots & \ddots & & & \vdots \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ \Phi_1(x_N) & \Phi_2(x_N) & \dots & \dots & \dots & \Phi_{M_w}(x_N) \end{pmatrix} \quad 4.42$$

On peut l'écrire de la façon suivante :

$$P = (P_1 \ P_2 \ \dots \ P_{M_w}) \quad 4.43$$

Avec

$$P_i = (\phi_i(x_1) \ \phi_i(x_2) \ \dots \ \phi_i(x_N))^T \quad \text{avec } i=1, \dots, M_w \quad 4.44$$

Les vecteurs p_i sont généralement linéairement indépendants (car $N \gg M_w$) et non orthogonaux. Les vecteurs p_i engendrent donc un sous-espace vectoriel de dimension M_w .

On estime que ces M_w vecteurs sont suffisants pour expliquer la sortie du réseau avec une précision satisfaisante. En d'autres termes, la projection du vecteur des sorties du processus Y_p dans cet espace correspond à une modélisation satisfaisante. La procédure de sélection consiste, en premier lieu, à classer les entrées par ordre de "pertinence" décroissante. Pour cela, on détermine, à chaque étape, l'ondelette qui a la plus grande projection sur la partie

du vecteur des sorties qui n'est pas expliquée par les entrées précédemment classées. La figure 4.7 propose une interprétation géométrique de cette procédure pour un exemple de dimension 2 (les ondelettes sont représentées par des vecteurs).

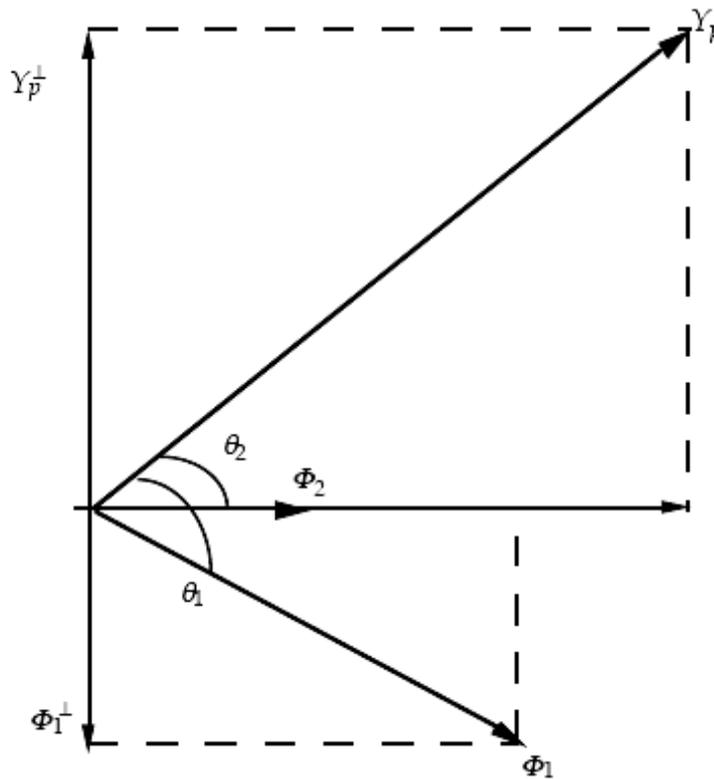


Figure 4.7: Interprétation géométrique de la sélection par orthogonalisation.

Sur cette figure, l'ondelette ϕ_2 est celle qui explique le mieux le vecteur des sorties Y_p puisque l'angle qu'elle fait avec ce vecteur est plus petit que celui entre ϕ_1 et Y_p . Elle est donc classée en premier rang par la procédure. Pour supprimer la partie de Y_p expliquée par ϕ_2 , on projette Y_p et les vecteurs correspondants aux ondelettes non encore classées (ici ϕ_1) dans l'espace orthogonal au vecteur que l'on vient de classer (ici ϕ_2). On a représenté ces projections par Y_p^\perp et ϕ_1^\perp .

Étant donné qu'on dispose des méthodes de construction de la bibliothèque d'ondelettes et de leur sélection, il reste à décrire les étapes de construction du réseau. Nous proposons le schéma suivant :

1. Effectuer l'apprentissage du réseau contenant uniquement les termes directs 'affine' (la solution est celle des moindres carrés).
2. Dédire une nouvelle séquence d'apprentissage dont les sorties sont définies comme les erreurs du réseau "affine". Cette séquence décrit donc la partie non modélisée par le réseau constitué par les termes directs.
3. Sélectionner un nombre N_w d'ondelettes de la bibliothèque préalablement construite, sur la base de la nouvelle séquence d'apprentissage.
4. Effectuer l'apprentissage du réseau complet, avec la séquence initiale, en ajustant les pondérations des ondelettes et les termes directs.

Lors du second apprentissage, on réajuste les termes directs pour ne pas aboutir à une solution sous-optimale. Une comparaison de leurs valeurs avant et après le second apprentissage montre qu'ils ne sont généralement pas modifiés.

Dans le cas où la taille du réseau n'est pas une contrainte, le nombre d'ondelettes N_w peut être augmenté tant qu'on n'observe pas de surajustement. En revanche, si l'on cherche à atteindre une performance donnée, on augmente N_w jusqu'à ce que cette performance soit atteinte. Dans les deux cas, la recherche de N_w se fait selon un processus itératif. La figure suivante illustre le schéma d'application de la procédure :

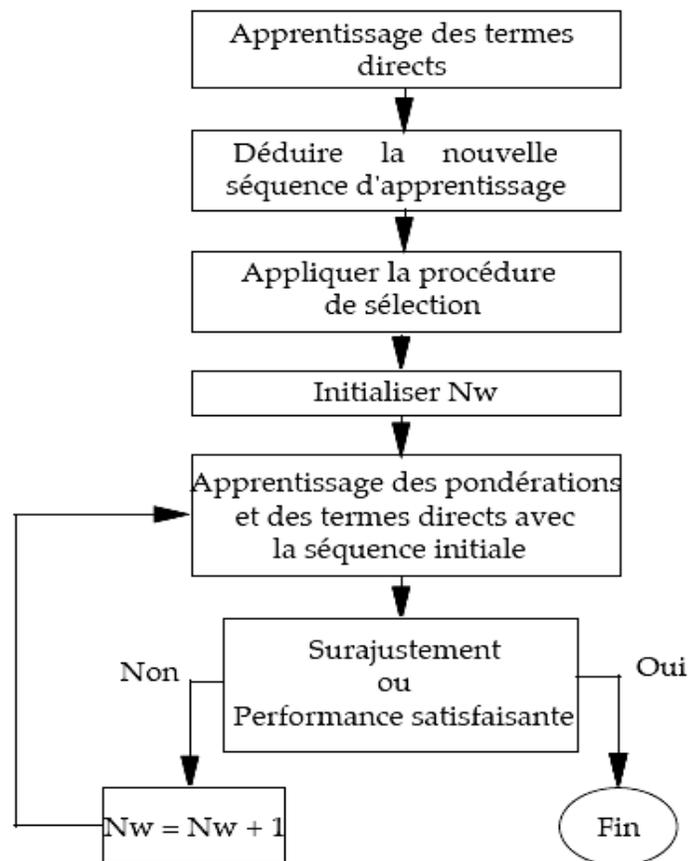


Figure 4.8: Schéma d'application de la procédure.

4.8 Conclusion

Ce chapitre présente une nouvelle technique basée sur les réseaux d'ondelettes qui peuvent faciliter la procédure de classification. Car ils correspondent à des réseaux de neurones qui ont des fonctions d'activation ressemblant aux ondelettes (dilatations et translations). Nous avons présenté une procédure de construction de réseaux d'ondelettes fondés sur la transformée discrète. Les réseaux obtenus sont plus volumineux, pour une même EQMP (l'erreur quadratique moyenne de performance), que ceux fondés sur la transformée continue, qui utilisent les techniques d'optimisation non linéaire. De plus, pour des modèles de trois entrées ou plus, le nombre d'éléments dans la bibliothèque devient très grand. En revanche, cette approche présente l'avantage d'utiliser les spécificités de la famille de fonctions (bibliothèque) pour une construction de réseaux à moindre coût (temps de calcul très inférieur à celui d'un apprentissage utilisant une technique d'optimisation non linéaire). Cette approche est donc conseillée dans le cas où l'on désire construire un modèle rapidement sans avoir recours à plusieurs apprentissages. C'est cette méthode que nous utiliserons pour classifier les EMG utérins.

Chapitre 5

Application - Adaptation aux signaux réels

5.1 Introduction

Ce chapitre est consacré à la présentation des résultats obtenus après avoir appliqué les algorithmes et les méthodes développées sur les signaux réels enregistrés.

Dans la première partie, nous allons présenter la méthode d'acquisition des signaux EMG utérins, et les caractéristiques de la population.

La deuxième partie présente l'ensemble de la base de données de signaux réels. Cette base peut être analysée à partir de deux groupes des signaux : tout d'abord, les signaux possédant le même terme à l'enregistrement (nombre de Semaines d'Aménorrhée à l'Enregistrement, SAR) mais avec des termes différents à l'accouchement (nombre de Semaines d'Aménorrhée à l'accouchement, SAA), regroupement qui permet de tester le pouvoir de classification en termes d'accouchement prématuré ou à terme ; ensuite, en regroupant les signaux possédant le même terme à l'accouchement (SAA), mais des termes différents à l'enregistrement (SAR). Ce regroupement permet de tester s'il est possible de caractériser le délai avant l'accouchement.

La troisième partie dans ce chapitre est consacrée aux méthodes de classification non supervisée et de paramétrage des signaux par l'utilisation de la décomposition en ondelettes. Le choix des ondelettes et des échelles sont discutés dans ce chapitre. Nous décrivons ainsi l'utilisation de la méthode basée sur le test de Fisher (Unsupervised Statistical Classification Method), ainsi que l'utilisation des réseaux de neurones compétitifs (Competitif Neural Network Method, CNNM).

Finalement, nous présenterons l'utilisation de modèle AR pour générer une basée des signaux simulés afin d'augmenter notre base de données. Puis nous décrivons les deux méthodes de classification supervisées : nouvelle méthode de classification basée sur les réseaux d'ondelettes (Wavelet network, Wavnet), réseaux à fonctions radiales (radial basis function, RBF). Une discussion sur tous les résultats obtenus sera enfin proposée.

5.2 Acquisition des signaux EMG utérins

L'enregistrement des signaux EMG a été effectué au Centre de Gynécologie et d'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 : grossesse monofaetale, patientes hospitalisées pour contractilité utérine excessive.

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.

Deux électrodes sont disposées à mi-hauteur de l'axe médian de l'utérus, entre le fond utérin et la symphyse pubienne.

Pendant la phase d'enregistrement, il est demandé à la femme de noter, à l'aide d'un marqueur d'événement manuel, tout mouvement de bébé ou toute contraction ressentis. Les marqueurs d'événements sont enregistrés en parallèle au 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 MAPs par électromyographie utérine [vois00].

5.3 Population

Les signaux réels dont nous disposons ont été enregistrés sur des femmes à plusieurs termes de grossesse et au cours de l'accouchement.

Les signaux enregistrés sur 25 femmes sont divisés en 5 groupes (tableau 5.1) selon les deux principaux critères suivants :

1. Semaine d'Aménorrhée à l'Enregistrement (SAR).
2. Semaine d'Aménorrhée à l'Accouchement (SAA)

Groupe	Nb placenta Antérieur/ postérieur	Nb de contractions/ nb de femmes	SAR ($\bar{M} \pm SD$)	SAA ($\bar{M} \pm SD$)
G1	4 / 2	22 / 6	29±0.76	33±0.65
G2	3 / 1	20 / 4	29±0.54	31±0.59
G3	2 / 4	22 / 6	29±0.43	36±0.45
G4	2 / 3	22 / 5	25±0.62	31±0.32
G5	2 / 2	16 / 4	27±0.45	31±0.34

Tableau 5.1. Caractéristiques de la population.

5.4 Classification Non supervisée

5.4.1 Méthodes choisies

Deux méthodes de classification sont utilisées : la première est basée sur le test de Fisher (USCM, Unsupervised Statistical Classification Method), cette nouvelle méthode a été développée dans le cadre de ce travail. La deuxième est une méthode basée sur les réseaux de neurones compétitifs (CNNM, Competitive Neural Network Method).

Les paramètres d'entrée des ces deux méthodes de classification sont calculés à partir des détails, après décomposition en ondelettes (variances sur les détails).

5.4.2 Paramétrage à partir de décomposition par ondelettes discrètes

Extraction des paramètres

Dans un premier temps, nous allons décomposer le signal en utilisant la décomposition en ondelettes ensuite nous analysons les détails sur quelques échelles. Les paramètres extraits de ces détails sont les variances sur les échelles [khal99].

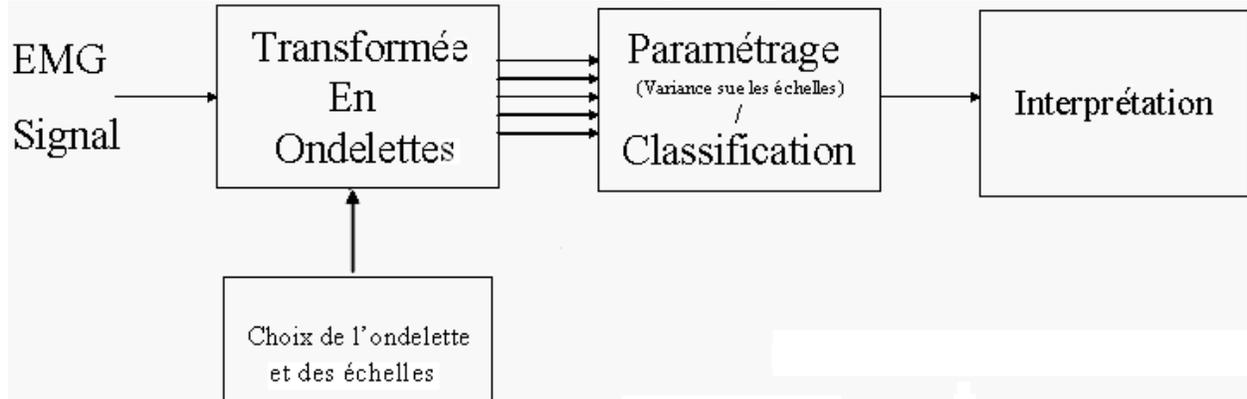


Figure 5.1. Procédure décomposition / classification.

Nous obtenons donc, pour chaque contraction un vecteur de variances qui correspond aux niveaux d'échelles (pour une classification en 5 niveaux, nous obtenons un vecteur à 5 variables pour chaque contraction).

Choix de l'ondelette

Le choix de l'ondelette analysante est un problème critique dans la théorie des ondelettes. Pour extraire des événements spécifiques dans un signal, le choix de l'ondelette doit être tel que la forme de l'ondelette soit adaptée à la forme de l'événement à détecter.

Dans les applications biomédicales, Li [li95] a utilisé une ondelette Spline quadratique qui lui permet de distinguer facilement les caractéristiques du bruit et de la ligne de base dans le signal électrocardiographique ECG. C'est une ondelette à support compact et à phase linéaire. Shenhaji [shen95] a utilisé une ondelette complexe pour la détection des ruptures dans l'ECG.

Dans une application de détection dans les signaux EMG utérins, l'ondelette de 'Symlet' (sym5) semble donner une meilleure possibilité de distinction [khal99]. Dans notre application, nous avons testé plusieurs ondelettes. Nous avons finalement utilisé la même ondelette de Symlet, 'sym5'.

Choix des échelles

Une fois choisie l'ondelette, il faut déterminer le nombre de niveaux d'échelles le plus approprié avant d'appliquer le paramétrage à chacune des échelles. Chaque niveau d'échelle présentant un contenu fréquentiel différent (du plus haut vers le plus bas en augmentant le nombre de niveaux). Il faut donc regarder la bande de fréquence qui correspond à chaque niveau d'échelle et l'énergie qu'elle contient afin de choisir les niveaux les plus convenables [abry97], [krim95].

Nous constatons que pour les EMG utérins dont le contenu fréquentiel est compris entre 0,2-3 Hz, nous pouvons nous limiter aux échelles 1 à 5 (figure 5.2). En effet, ces 5 niveaux regroupent plus de 96% de l'énergie du signal et couvrent la bande de fréquence d'intérêt.

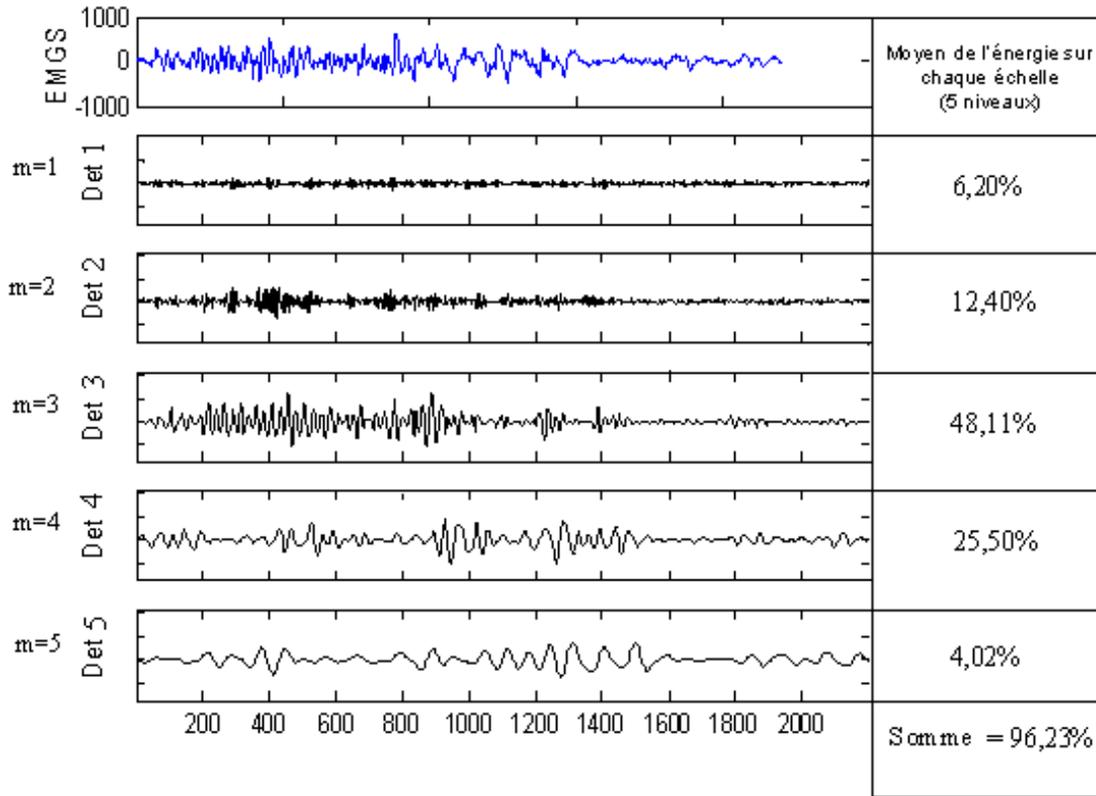


Figure 5.2. Décomposition par Ondelette 'Sym5' sur 5 niveaux d'échelle. La colonne de droite montre le pourcentage d'énergie dans les sous signaux de détails. On remarque que 5 sous signaux contiennent 96% de l'énergie totale du signal.

5.4.3 Applications aux signaux synthétiques

Génération des signaux synthétiques

Afin de pouvoir tester les méthodes, la décomposition et la classification sont appliquées initialement sur des signaux synthétiques. Des segments successifs sont générés en utilisant le filtrage d'un bruit gaussien (nous supposons que les signaux EMG utérin suivent des lois gaussiennes) avec un filtre passe bande ayant des fréquences centrales différentes mais de même largeur de bande (figure 5.3), ce qui permet de simuler des différences de contenu fréquentiel, phénomène observé entre les contractions de grossesses et celles d'accouchement [diab04a].

Soient S_i (S_j) deux segments correspondants à un filtrage du bruit blanc par un filtre passe bande ayant deux fréquences centrales f_i et f_j .

La résolution de la classification implique l'analyse de la distance entre S_i et S_j .

$$\delta_{ij} = \frac{\Delta F_{ij}}{BP} = \frac{f_i - f_j}{BP}. \quad 5.1$$

BP: bande passante; ΔF_{ij} : différence entre les fréquences centrales.

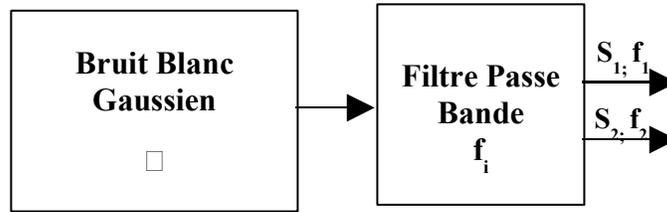


Figure 5.3. Génération des segments à l'aide des signaux synthétiques.

Nous utilisons une BP fixe pour les signaux synthétiques. Pour cette raison la variation du δ_{ij} est liée directement à la variation de ΔF_{ij} .

La distance entre les classes est évidemment augmentée quand ΔF_{ij} augmente [diab06a]. Nous avons utilisé 20 segments pour chaque classe.

Résultats sur les signaux synthétiques

Après avoir généré les segments, les paramètres d'ondelettes sont utilisés (variance sur les échelles) en entrée de la méthode de classification non supervisée basée sur le test de Fisher (USCM, paragraphe 2.7) et de la méthode basée sur les réseaux de neurones compétitifs (CNNM).

La classification a été faite d'une façon non supervisée (séparation en deux classes). Ensuite les résultats ont été comparés aux groupes réels pour interpréter les résultats.

Les résultats présentés figures 5.4 et 5.5 utilisent deux classes des signaux, avec respectivement ΔF_{ij} égal à 0.1 et à 0.5 (fréquence normalisée $f_{\max}=1$), avec la méthode USCM.

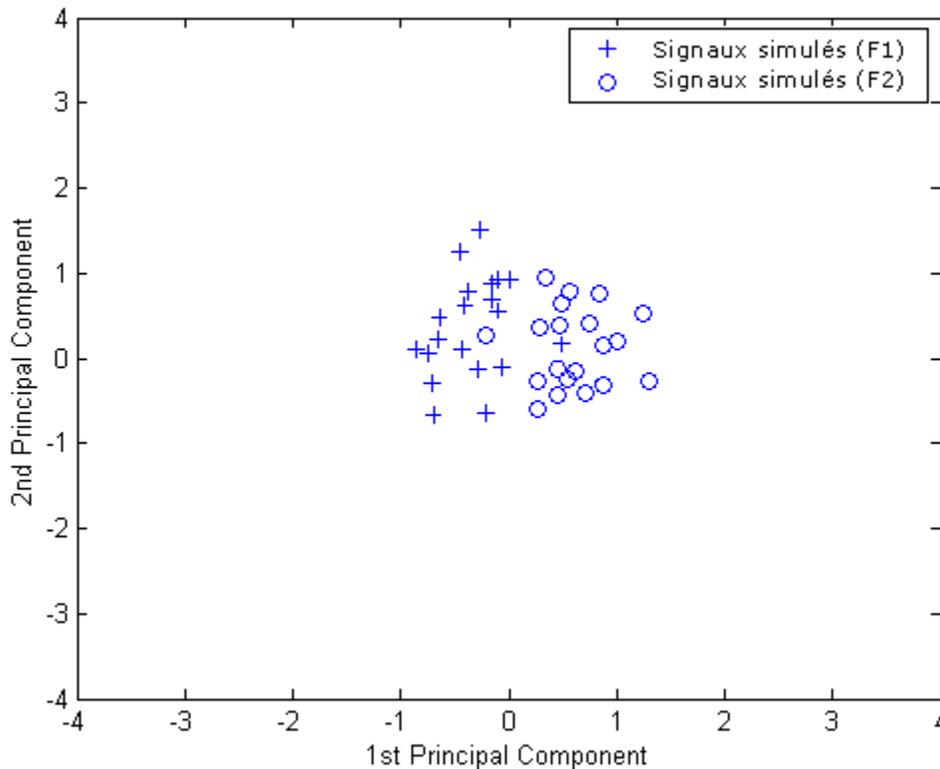


Figure 5.4. Projection par Analyse en Composante Principale : Classification des signaux synthétiques avec 2 fréquences centrales, $\Delta F_{ij} = 0.1$ (méthode USCM). Ondelette : 'sym5', échelle : de 1 à 5. Variance projetée : 76%.

	F ₁	F ₂
F ₁	19	1
F ₂	1	19

Tableau 5.2. Matrice de confusion : résultat de la classification sur les signaux simulés avec 2 fréquences centrales, $\Delta F_{ij} = 0.1$ (méthode USCM).

	F ₁	F ₂
F ₁	19	1
F ₂	2	18

Tableau 5.3. Matrice de confusion : résultat de la classification sur les signaux simulés avec 2 fréquences centrales, $\Delta F_{ij} = 0.1$ (méthode CNNM).

Sur la figure 5.4, nous pouvons noter une séparation faible. La matrice de confusion correspondante est montrée dans le tableau 5.2 (USCM, erreur de classification égale à 5%). Le tableau 5.3 nous montre la matrice de confusion obtenue par la méthode de réseau des neurones compétitif (CNNM, erreur de classification égale à 7.5%). L'erreur est importante car la distance entre les classes est petite (0.1).

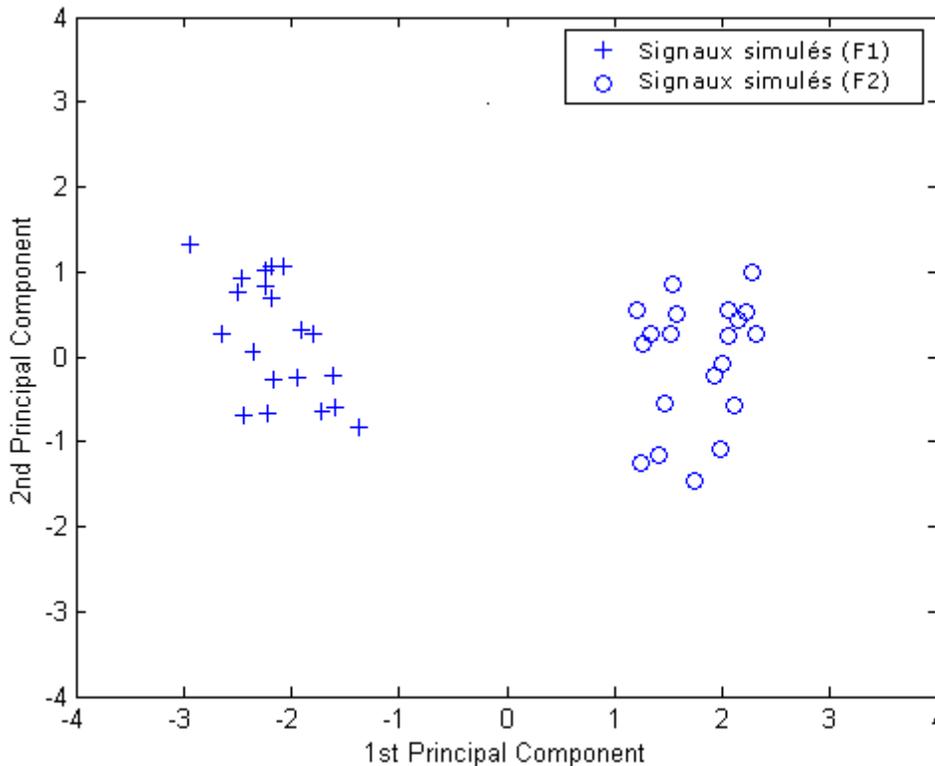


Figure 5.5. Projection par Analyse en Composante Principale : classification des signaux synthétiques avec 2 fréquences centrales, $\Delta F_{ij} = 0.5$ (méthode USCM). Ondelette : 'sym5', échelle : de 1 à 5. Variance projetée : 85%.

Dans la figure 5.5 la séparation est complète entre les deux classes. La matrice de confusion correspondante à la méthode d'USCM est montrée dans le tableau 5.4 (erreur de la classification égale à 0%). Le tableau 5.5 nous montre les résultats obtenus par CNNM (erreur de classification égale à 0%). Aucune erreur ne se produit parce que la distance entre les classes est très grande (0.5).

	F ₁	F ₂
F ₁	20	0
F ₂	0	20

Tableau 5.4. Matrice de confusion : résultat de la classification sur les signaux simulés avec 2 fréquences centrales, $\Delta F_{ij} = 0.5$ (méthode USCM).

	F ₁	F ₂
F ₁	20	0
F ₂	0	20

Tableau 5.5. Matrice de confusion : résultat de la classification sur les signaux simulés avec 2 fréquences centrales, $\Delta F_{ij} = 0.5$ (méthode CNNM).

Ces deux méthodes sont donc capables de mettre en évidence des différences dans le contenu spectral des signaux. Elles sont de plus sensibles à la distance entre les classes.

5.4.4 Application aux signaux réels

Les méthodes de classification ensuite ont été utilisées sur des signaux réels. Pour cela, nous les avons appliquées pour classifier les groupes de contractions ayant le même SAR (Semaine d'Aménorrhée d'Enregistrement) mais ayant différents SAA (Semaine d'aménorrhée à l'Accouchement). Puis pour classifier des signaux ayant même SAA et différents SAR.

Les paramètres utilisés sont les variances calculées après décomposition par ondelettes sur 5 niveaux d'échelle, en utilisant l'ondelette Symlet ('sym5') [diab04b], [diab05].

Résultats sur les groupes ayant même SAR et différents SAA

Afin d'examiner la possibilité de détecter le terme à l'accouchement, à une date d'enregistrement donnée, nous avons d'abord utilisé les groupes qui ont les mêmes SAR et différents SAA. Le premier groupe G1 correspond aux signaux enregistrés à 29 SAR dont les femmes ont accouché à 33 SAA. Le deuxième groupe G2 correspond aux signaux enregistrés à 29 SAR dont les femmes ont accouché à 31 SAA. Nous pouvons noter sur la figure 5.6 une séparation partielle entre les deux groupes. Les résultats sont également présentés sous forme de matrices de confusion (tableaux 5.6 et 5.7). L'erreur de classification calculée est 9.5% (9.5% de contractions sont mal classifiés) en appliquant la méthode de classification statistique non-supervisée (USCM), et de 14.2% en appliquant les réseaux de neurones compétitifs (CNNM).

Après classification non supervisée (clustering en deux groupes) les résultats sont comparés avec l'aide d'un expert pour valider les résultats.

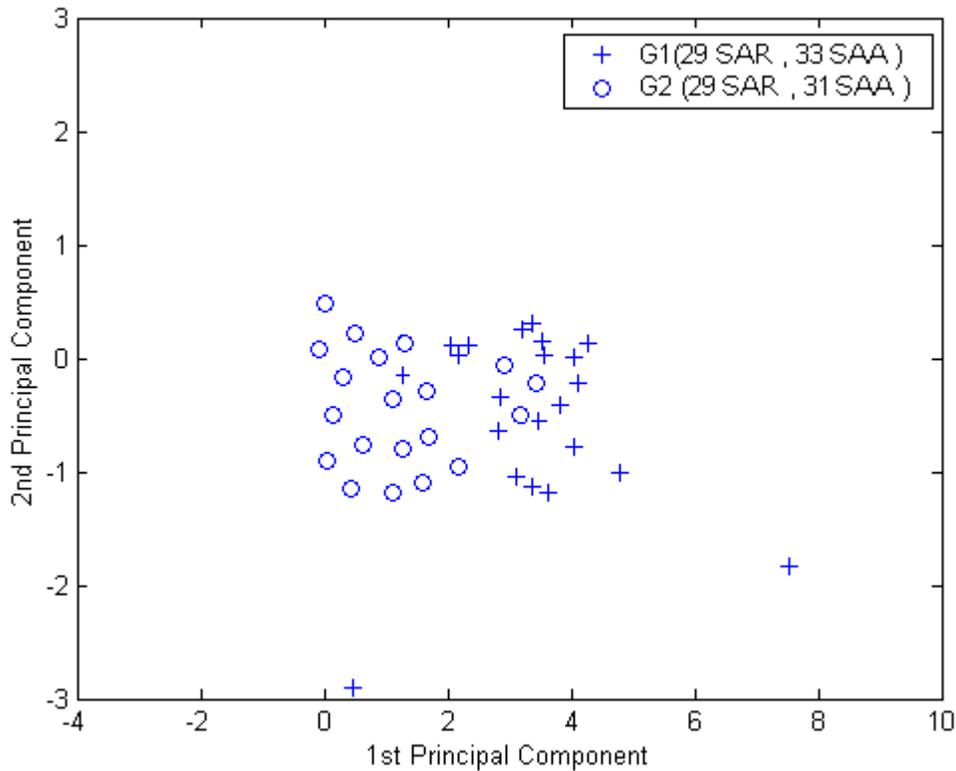


Figure 5.6. Projection par Analyse en Composante Principale : classification de G1 et G2 possédant le même SAR et une petite différence de SAA (méthode USCM). Ondelette : 'sym5', échelle : de 1 à 5. Variance projetée : 72%.

L'erreur dans cette représentation est importante et la séparation entre les deux groupes n'est pas bonne parce que les deux groupes de mêmes SAR ont des valeurs très proches de SAA.

	G ₁	G ₂
G ₁	21	1
G ₂	3	17

	G ₁	G ₂
G ₁	19	3
G ₂	3	17

Tableau 5.6. Matrice de confusion : résultat de la classification sur G1 et G2 ayant le même SAR et une petite différence en SAA (méthode USCM), temps de calcul 3s.

Tableau 5.7. Matrice de confusion : résultat de la classification sur G1 et G2 ayant le même SAR et une petite différence en SAA (méthode CNNM), temps de calcul 17s.

La figure 5.7, présente les résultats de la classification de deux autres groupes. Le premier groupe est G3 et correspond aux signaux enregistrés à 29 SAR dont les femmes ont accouché à 36 SAA. Le deuxième groupe G2 correspond aux signaux enregistrés à 29 SAR et dont les femmes ont accouché à 31 SAA. Sur la figure 5.8 nous pouvons noter une séparation totale entre les deux groupes. Les erreurs de classification sont égales à 4.7% (calculé à partir du tableau 5.8) en appliquant l'USCM et à 9.5% (calculé à partir du tableau 5.9) en appliquant le CNNM.

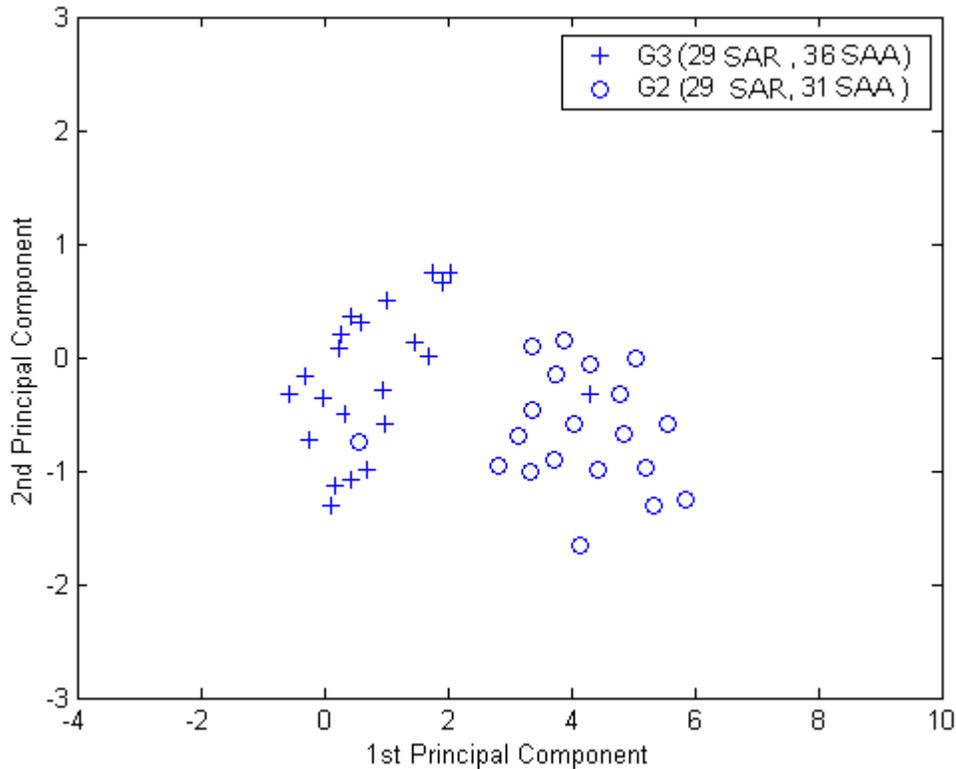


Figure 5.7. Projection par Analyse en Composante Principale : classification de G3 et G2 ayant le même SAR et grande différence dans SAA (méthode USCM). Ondelette : 'sym5', échelle : de 1 à 5. Variance projetée : 76%.

L'erreur dans cette représentation est faible et la séparation est claire parce que les deux groupes de même SAR ont une plus grande différence entre les valeurs de SAA.

	G ₃	G ₂
G ₃	21	1
G ₂	1	19

Tableau 5.8. Matrice de confusion : résultat de la classification sur G3 et G2 ayant la même SAR et une grande différence en SAA (méthode USCM) temps de calcul 3s.

	G ₃	G ₂
G ₃	20	2
G ₂	2	18

Tableau 5.9. Matrice de confusion : résultat de la classification sur G3 et G2 ayant la même SAR et une grande différence en SAA (méthode CNNM) temps de calcul 13s.

Résultats sur les groupes ayant même SAA et différents SAR

La deuxième étape était d'examiner la possibilité de différencier des contractions ayant le même terme d'accouchement (SAA) mais différents termes à l'enregistrement (SAR). Cette approche permet de savoir s'il est possible distinguer de différentes évolutions au cours de la grossesse avant une date d'accouchement donnée. Deux groupes sont utilisés, le premier

G2 correspond aux signaux enregistrés à 29 SAR et correspond aux femmes qui ont accouché à 31 SAA. Le deuxième groupe G5 correspond aux signaux enregistrés à 27 SAR pour des femmes qui ont également accouché à 31 SAA.

Nous pouvons voir figure 5.8 une séparation partielle entre les deux groupes. L'erreur associée est 13.8 % (calculés à partir du tableau 5.10) en appliquant l'USCM et 16.6 % (calculés à partir du tableau 5.11) en appliquant le CNNM. Les deux groupes ne sont pas bien séparés parce que les différences entre les SAR des deux groupes sont très étroites.

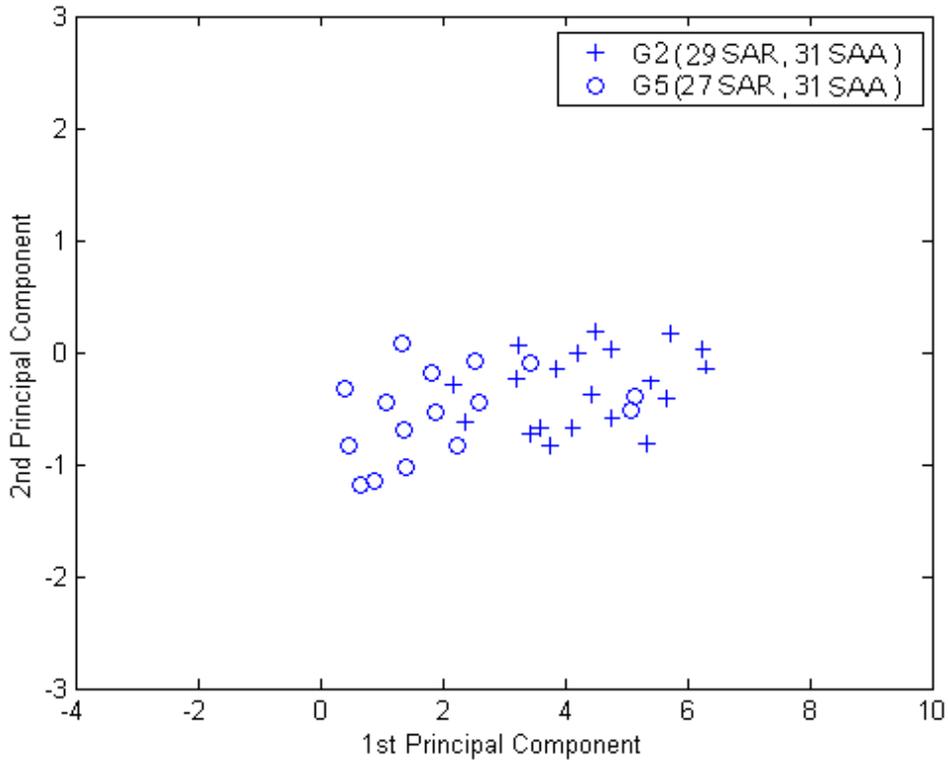


Figure 5.8. Projection par Analyse en composante principale : classification de G2 et G5 ayant le même SAA et petite différence dans SAA (méthode USCM). Ondelette : 'sym5', échelle : de 1 à 5. Variance projetée : 74%.

	G₂	G₅
G₂	18	2
G₅	3	13

Tableau 5.10. Matrice de confusion : résultat de la classification sur G2 et G5 ayant le même SAA et une petite différence en SAR (méthode USCM) temps de calcul 2.5s.

	G₂	G₅
G₂	18	2
G₅	4	12

Tableau 5.11. Matrice de confusion : résultat de la classification sur G2 et G5 ayant le même SAA et une petite différence en SAR (méthode CNNM), temps de calcul 20s.

Deux autres groupes ont été étudiés, le premier G4 correspond aux signaux enregistrés à 25 SAR pour des femmes qui ont accouché à 31 SAA et le deuxième groupe G2 correspond aux signaux enregistrés à 29 SAR pour des femmes ayant accouché également à 31 SAA. La figure 5.9 montre une séparation presque totale entre les deux groupes. Les tableaux 5.12 et 5.13 montrent les matrices de confusions obtenues pour les méthodes USCM et CNNM avec une erreur respectivement de 2.3% et 7.1%.

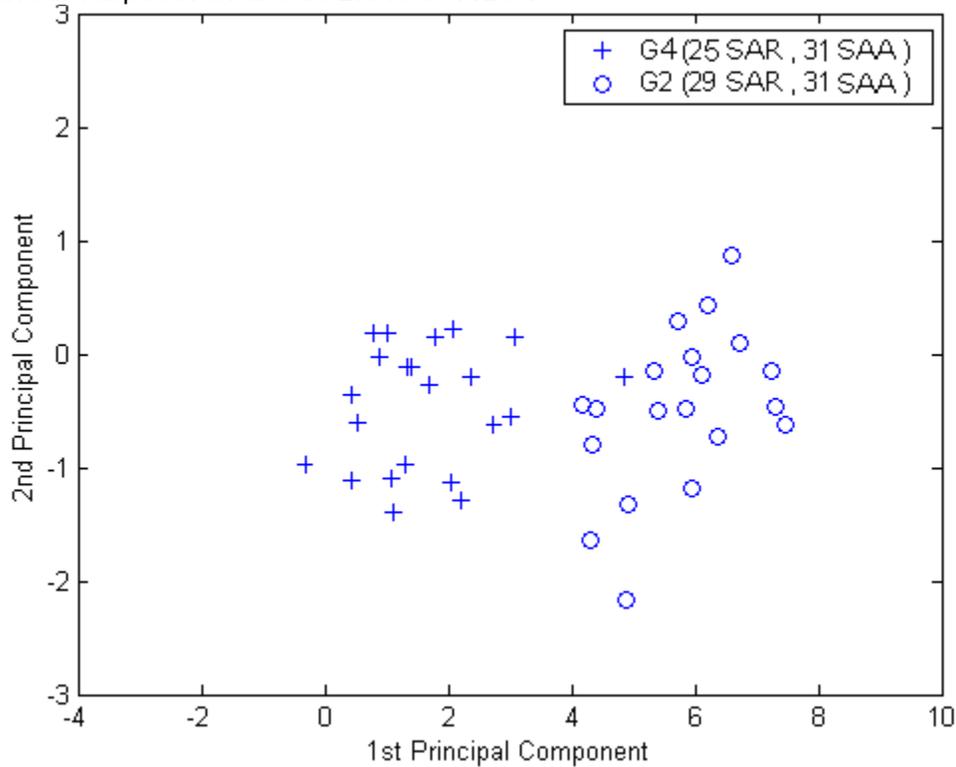


Figure 5.9. Projection par Analyse en Composante Principale : Classification de G4 et G2 ayant le même SAA et grande différence dans SAR (méthode USCM). Variance projetée : 79%.

	G ₄	G ₂
G ₄	21	1
G ₂	0	20

Tableau 5.12. Matrice de confusion : résultat de la classification sur G4 et G2 ayant le même SAA et une grande différence en SAR (méthode USCM), temps de calcul 3s.

	G ₄	G ₂
G ₄	20	2
G ₂	1	19

Tableau 5.13. Matrice de confusion : résultat de la classification sur G4 et G2 ayant le même SAA et une grande différence en SAR (méthode CNNM) temps de calcul 9s.

5.4.5 Conclusion sur les méthodes non supervisées

Dans cette partie on a présenté les résultats de classification des méthodes non-supervisées, qui ne nécessitent ni apprentissage, ni connaissance a priori sur les signaux de la base de données. D'après les résultats dans le tableau 5.12, toutes les contractions sont presque classifiées. Ce qui montre que plus la différence entre un terme à l'accouchement donné et le terme à l'enregistrement augmente, plus les signaux sont différents. Il semble donc qu'il y ait une évolution du contenu fréquentiel du signal au cours de la grossesse. Et on peut noter dans cette partie que la méthode USCM donne de meilleurs résultats que les réseaux de neurones compétitifs (CNNM). De plus, si on compare le temps de calcul entre les deux méthodes, l'USCM est plus rapide que la méthode CNNM.

5.5 Classification Supervisée utilisant le Réseau à Fonctions Radiales et le Réseau d'Ondelettes

5.5.1 Méthodes choisies

Pour les méthodes de classification appliquées ici sur les signaux EMG utérin, particulièrement le réseau d'ondelettes, les coefficients d'ondelettes sont les fonctions d'activation de la couche cachée. La caractérisation par ondelettes est donc incluse dans la méthode de classification elle-même.

Cette nouvelle technique basée sur les réseaux d'ondelettes nous facilite la procédure de classification car on peut utiliser directement un réseau d'ondelette, qui est un réseau de neurones normal et qui a des fonctions d'activation qui ressemblent aux ondelettes (dilatations et translations).

Ce réseau est très utilisé pour la régression. Nous l'utilisons ici pour la classification.

Nous avons aussi utilisé les réseaux de neurones à fonction radiales comme autre méthode de classification. Cette méthode est supervisée et c'est la plus proche de la structure des réseaux d'ondelettes.

En ce qui concerne l'ensemble d'apprentissage, nous présenterons dans cette section les méthodes d'estimation AR et le choix de l'ordre du modèle, ainsi que la technique utilisée pour augmenter l'ensemble d'apprentissage.

Nous utiliserons dans cette partie uniquement les paramètres extraits de la densité spectrale de puissance.

5.5.2- Paramétrage à partir de la densité spectrale de puissance

Dans cette section, nous utilisons les déciles comme paramètres extraits de la densité spectrale de puissance (figure 5.10), calculée à partir du périodogramme de Welch, qui produit une estimation beaucoup plus robuste que le modèle AR lorsque le niveau de bruit est important [kunt91], [kay88], [lema96]. La figure 5.10 présente les déciles calculés à partir d'un spectre EMG.

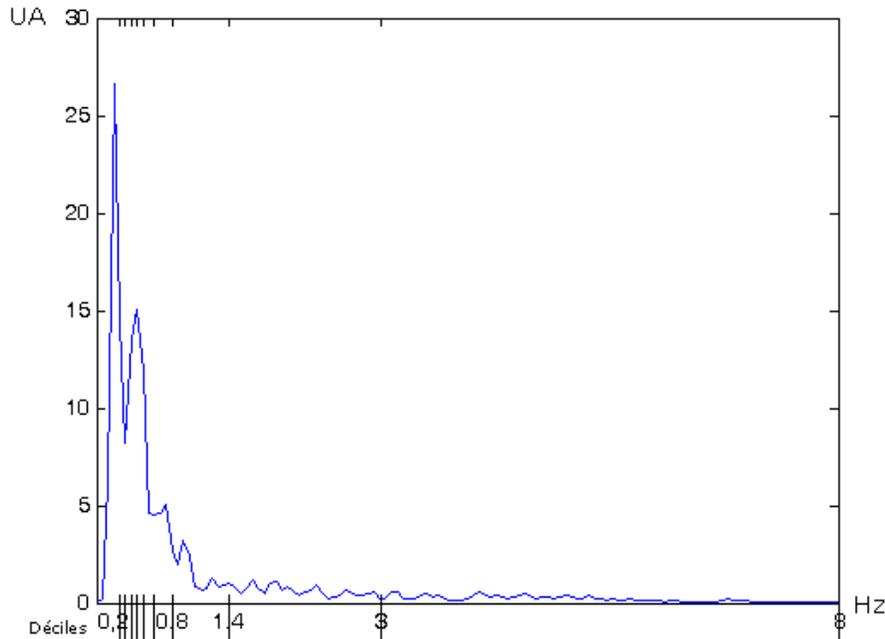


Figure 5.10. Extraction des déciles à partir de la PSD d'un EMG (Axe verticale : U.A. Unités Arbitraires, Axe horizontale : Hz).

5.5.3 Problème d'apprentissage et augmentation de la base de données

Comme les méthodes de classification supervisées nécessitent un grand nombre d'individus dans l'ensemble d'apprentissage et comme les données réelles dont nous disposons sont limitées, nous avons généré des segments d'EMG utérins simulés après avoir modélisé chaque groupe de signaux réels par des modèles AR. Cette génération des signaux simulés a été faite en utilisant un bruit blanc en entrée d'un filtre AR qui modélise l'EMG utérin.

Avant d'appliquer une méthode de filtrage AR, il faut choisir l'ordre et la méthode d'estimation.

Choix de la méthode d'estimation du modèle AR

Le choix de la méthode d'estimation est un point très important, qui doit être lié à l'utilisation ultérieure des paramètres estimés. La comparaison entre les méthodes d'estimation va être faite ici sur la base des performances obtenues par l'algorithme de classification associé aux modèles AR estimés.

Ce choix sera justifié selon les résultats de la classification : distance entre les classes et regroupement des individus qui appartiennent à la même classe (distance intra classe tableau 5.14) [diab06b].

Choix de l'ordre du modèle et longueur bruit blanc gaussien

Dans le cas de signaux réels, l'ordre du modèle AR utilisé pour les modéliser n'est pas connu a priori. Plusieurs critères sont utilisés pour la sélection de l'ordre, ces critères étant principalement fondés sur l'évolution de l'erreur de prédiction [kay88], [marp87].

Quand les signaux réels sont représentés par un modèle AR, il n'y a pas un ordre unique quel que soit le type de signal. En effet l'ordre varie en fonction du protocole expérimental,

des conditions de recueil des signaux et de leur utilisation. Par exemple, dans [heff88] l'ordre choisi est 4 (pour les signaux EMG), alors que d'autres comme [grau89] utilisent un ordre 6. Dans le cas d'estimation spectrale, l'ordre utilisé par [seid87] est de 15, dans [falo02] un ordre de 20 a été utilisé, alors que dans [coat83] un ordre de 8 est utilisé pour le même problème d'estimation spectrale.

Dans notre cas, un ordre de 16 a été utilisé [diab06b] associé à l'algorithme de covariance. Ce choix est justifié par les résultats de l'ACP, distance entre les classes et le regroupement des individus qui appartiennent au même groupe (tableau 5.14). Ce choix est basé sur le calcul de la variance intraclasse (σ_1 et σ_2) et interclasse (d) après projection ACP. Nous avons obtenu le tableaux suivant :

Order	Distance	Yule-Walker (Y.W)	Burg (B)	Covariance (C)	Covariance modifiée (C.M)
10	σ_1	0.63	0.51	0.45	0.64
	σ_2	0.72	0.62	0.55	0.72
	d	2.1	2.35	2.4	2.2
16	σ_1	0.52	0.53	0.45	0.52
	σ_2	0.64	0.85	0.57	0.62
	d	2.81	2.94	3.52	3.14
20	σ_1	0.78	0.71	0.56	0.56
	σ_2	0.83	0.86	0.59	0.55
	d	3.01	2.73	3.12	2.95

Tableau 5.14. Comparaison entre des méthodes d'estimation (YW, B, C, CM) et de l'ordre de chaque méthode (10, 16, 20) pour les signaux EMG utérins, σ_1 : distance moyenne entre les points et le centre du classe 1; σ_2 : distance moyenne entre les points et le centre du classe 2; d : distance entre le centre du classe 1 et le centre du classe 2 .

Nous avons aussi testé l'influence de la longueur du bruit blanc en entrée du modèle AR (tableau 5.15).

Méthode / ordre	Distance	Longueur de bruit blanc gaussien		
		N=500	N=2000	N=4000
Covariance Ordre:16	σ_1	0.53	0.55	0.69
	σ_2	0.62	0.66	0.75
	d	1.85	3.74	3.05

Tableau 5.15. Influence de la longueur du bruit blanc gaussien (100, 500, 2000) à l'entrée du modèle AR, sur la séparabilité des classes σ_1 : distance moyenne entre les points et le centre du classe 1; σ_2 : distance moyenne entre les points et le centre du classe 2; d : distance entre le centre du classe 1 et le centre du classe 2 .

Nous avons donc retenu la méthode Covariance avec un ordre 16 et un nombre de points du signal d'entrée (Bruit Blanc) de 2000.

Génération des modèles APs et ATs

Dans notre étude un modèle AR a été utilisé pour estimer les coefficients AR (a_i) des deux types des signaux d'EMG : EMG d'accouchement prématuré (APs), EMG accouchement à terme (ATs). Nous avons utilisé 10 contractions de chaque type (figure 5.11) pour avoir le modèle de base.

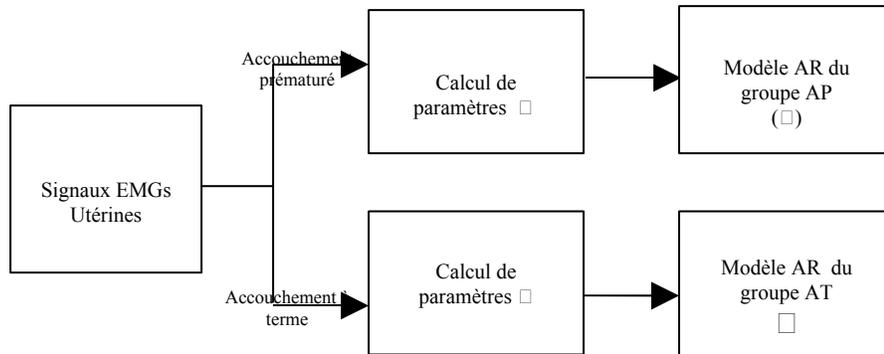


Figure 5.11. Création des modèles AR pour AP et AT.

Après obtention des paramètres AR pour chaque contraction, nous avons moyenné chaque ensemble de paramètres. Les paramètres moyens sont considérés comme le modèle de chaque groupe.

En utilisant les deux modèles AR de deux groupes APs et ATs, nous avons généré 100 signaux de chaque type qui simulent les deux types de contractions utérines : signaux EMGs d'accouchements prématurés simulés et signaux EMGs à terme simulés, en utilisant à l'entrée des modèles un bruit blanc gaussien (figure 5.12).

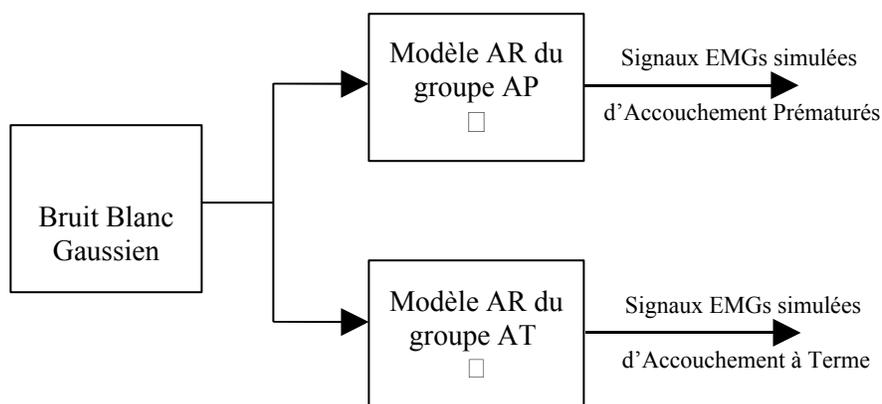


Figure 5.12. Génération des signaux EMGs AP simulés et signaux EMGs AT.

Ces signaux simulés pour chaque classe, sont ensuite utilisés pour faire l'apprentissage des deux méthodes utilisées : Wavnet et RBF.

Réseaux d'ondelettes (Wavnet)

Dans cette méthode nous avons utilisé les déciles extraits de la DSP d'une contraction EMG utérin comme paramètres de classification.

Les signaux simulés générés par le model AR (100 signaux de chaque type APs et ATs) sont utilisés pour la procédure d'apprentissage (section 4.7.2). Le nombre d'ondelettes utilisé a été choisie égal à 8, après comparaison des résultats obtenus avec 5, 8 et 10 ondelettes [chke07]. Nous avons travaillé avec l'ondelette de la dérivée première de la fonction gaussienne [ouss98].

Après l'apprentissage, nous avons présenté à l'entrée du réseau les déciles de deux groupes des contractions (AP, AT) qui n'avaient pas été utilisées lors de la génération de la base de données de signaux simulés.

Réseaux à fonction radiales (RBF)

Les paramètres utilisés dans les réseaux RBF sont les suivants :

Nombre d'unités cachées : 20

Fonction d'activation de la couche cachée : fonction gaussienne

Fonction de la couche de sortie : linéaire

Ces trois paramètres seront utilisés, dans le cadre de la classification d'accouchement des femmes avec le RBF. De même, l'apprentissage utilisera les signaux simulés générés par le model AR (100 signaux de chaque type APs et ATs) et les paramètres d'entrée sont les déciles de chaque contraction.

5.5.6 Application aux signaux réels

Classification utilisant les déciles de la DSP

Pour la classification des accouchements à termes (ATs) ou accouchements prématurés (APs), nous calculons la DSP de la contraction à classifier, puis les déciles extraites à partir de sa DSP sont appliqués à l'entrée de Wavnet ou RBF (figure 5.14).

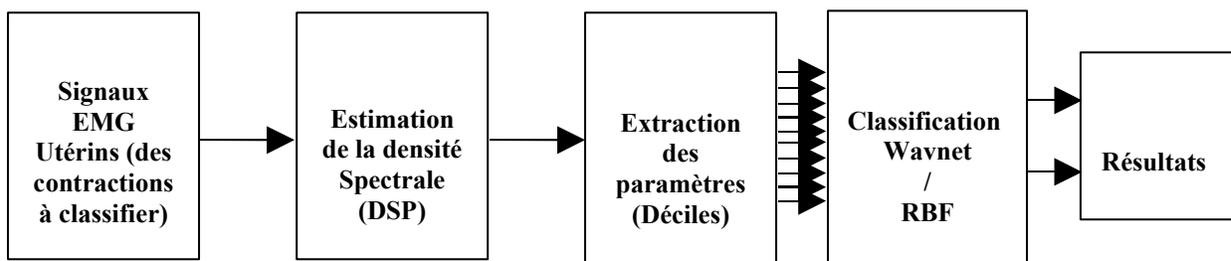


Figure 5.13. Schéma de classification de nouvelles données.

Nous utiliserons pour les tests les même 5 groupes de signaux que ceux utilisés avec les méthodes non supervisées (tableau 5.1).

Résultats sur les groupes ayant même SAR et différents SAA

Le premier groupe G1 correspond aux signaux enregistrés à 29 SAR et dont les femmes ont accouché à 33 SAA. Le deuxième groupe G2 correspond aux signaux enregistrés à 29 SAR et dont les femmes ont accouché à 31 SAA. Les résultats sont présentés comme matrice de

confusion (tableaux 5.16 et 5.17). L'erreur de classification calculée est de 7.1% (7.1% de contractions sont mal classifiés) pour la méthode de classification de réseaux d'ondelettes (wavnet), et de 9.5% pour le réseau à fonction radiale (RBF).

	G ₁	G ₂
G ₁	21	1
G ₂	2	18

Tableau 5.16. Matrice de confusion : résultat de la classification de G1 et G2 ayant le même SAR et une petite différence de SAA (méthode Wavnet), temps de calcul de 2min.

	G ₁	G ₂
G ₁	20	2
G ₂	2	18

Tableau 5.17. Matrice de confusion : résultat de la classification de G1 et G2 ayant le même SAR et une petite différence de SAA (méthode RBF), temps de calcul de 40 s.

En ce qui concerne la classification du groupe G3 qui correspond aux signaux enregistrés à 29 SAR et aux femmes qui ont accouché à 36 SAA, et du groupe G2 qui correspond aux signaux enregistrés à 29 SAR et aux femmes qui ont accouché à 31 SAA, l'erreur de classification obtenue est respectivement égale à 2.3% (calculé à partir de tableau 5.18) en appliquant Wavnet et 7.1% (calculé à partir de tableau 5.19) en appliquant le RBF. On observe une réduction de l'erreur pour les deux méthodes quand la différence entre terme à l'enregistrement et terme à l'accouchement augmente.

	G ₃	G ₂
G ₃	21	1
G ₂	0	20

Tableau 5.18. Matrice de confusion : résultat de la classification de G3 et G2 ayant le même SAR et une grande différence de SAA (méthode Wavnet), temps de calcul de 1min 50s.

	G ₃	G ₂
G ₃	21	1
G ₂	2	18

Tableau 5.19. Matrice de confusion : résultat de la classification de G3 et G2 ayant le même SAR et une petite différence de SAA (méthode RBF), temps de calcul de 30s.

Résultats sur les groupes ayant même SAA et différents SAR

Deux groupes sont utilisés, le premier G2 correspond aux signaux enregistrés à 29 SAR et correspond aux femmes qui ont accouché à 31 SAA. Le deuxième groupe G5 correspond aux signaux enregistrés à 27 SAR et aux femmes qui ont accouché également à 31 SAA. L'erreur associée est de 11.1% (calculés à partir de tableau 5.20) en appliquant Wavnet, et égale à 16.6% (calculés à partir de tableau 5.21) en appliquant RBF. Les deux groupes ne sont pas bien séparés parce que les SAR pour les deux groupes sont proches.

	G ₂	G ₅
G ₂	18	2
G ₅	2	14

Tableau 5.20. Matrice de confusion : résultat de la classification de G₂ et G₅ ayant le même SAA et une petite différence de SAR (méthode Wavnet), temps de calcul de 1min 70s.

	G ₂	G ₅
G ₂	18	2
G ₅	4	12

Tableau 5.21. Matrice de confusion : résultat de la classification de G₂ et G₅ ayant le même SAA et une petite différence de SAR (méthode RBF), temps de calcul de 45s.

Deux autres groupes sont étudiés, le premier G₄ correspond aux signaux enregistrés à 25 SAR et dont les femmes ont accouché à 31 SAA, et le deuxième groupe G₂ correspond aux signaux enregistrés à 29 SAR et dont les femmes ont aussi accouché à 31 SAA. Les tableaux 5.22 et 5.23 montrent les matrices de confusions pour le Wavnet et RBF avec une erreur de classification respectivement égale à 2.8% et 4.7%.

	G ₄	G ₂
G ₄	21	1
G ₂	0	20

Tableau 5.22. Matrice de confusion : résultat de la classification de G₄ et G₂ ayant le même SAA et une grande différence de SAR (méthode Wavnet), temps de calcul de 1min 15s.

	G ₄	G ₂
G ₄	21	1
G ₂	1	19

Tableau 5.23. Matrice de confusion : résultat de la classification de G₄ et G₂ ayant le même SAA et une grande différence de SAR (méthode RBF), temps de calcul de 25s.

5.6 Discussion

Dans ce chapitre quatre méthodes de classification ont été utilisées pour séparer différents groupes. Les deux premières méthodes sont des méthodes de classification non-supervisées : une méthode basée sur la classification statistique non supervisée et une autre méthode basée sur les réseaux de neurones compétitifs. Les deux dernières sont des méthodes de classification supervisée, une basée sur les réseaux d'ondelettes et une autre basée sur les réseaux à fonction radiales. Toutes ces méthodes sont appliquées sur des signaux EMGs utérins réels.

Pour ce qui concerne la classification non-supervisée, nous avons noté dans la première partie, que la méthode USCM donne de meilleurs résultats que les réseaux des neurones compétitifs (CNM). On note que pour les deux approches de classification (SAR fixe et SAA variable ou SAA fixe et SAR variable), la méthode de classification statistique donne des

erreurs de classification plus faible que les réseaux de neurones compétitifs. En comparant les temps de calcul des deux méthodes, nous observons que l'USCM est plus rapide que la méthode CNNM. L'avantage des méthodes de classification non supervisées vient du fait qu'on peut les utiliser si on n'a pas de connaissance a priori sur les signaux de base de donnée. Cependant des études a posteriori sont nécessaires pour valider le regroupement.

Pour ce qui concerne la classification supervisée, nous avons noté dans la deuxième partie (5.5), que la méthode basée sur les réseaux d'ondelettes (Wavnet) donne des meilleurs résultats que les réseaux à fonction radiales (RBF). Dans les deux approches de classification (SAR fixe et SAA variable ou SAA fixe et SAR variable) les réseaux d'ondelettes donnent une erreur de classification plus faible que celles donnée par les réseaux à fonction radiales. En comparant le temps de calcul entre les deux méthodes, le RBF est beaucoup plus rapide que le Wavnet. De plus la méthode Wavnet nécessite un temps important pour la phase d'initiation des paramètres. La méthode RBF est plus facile de mettre en œuvre et beaucoup plus rapide mais la méthode Wavnet est plus précise.

En comparant nos résultats avec les résultats de l'équipe de Garfield [mane07] nous avons obtenu des pourcentages plus élevés. En ce qui concerne les accouchements à terme, nous avons obtenu une erreur de classification de 4,54% avec l'USCM (tableau 5.8) et de 9,09% avec le CNNM (tableau 5.9), alors que l'équipe de Garfield avait obtenu une erreur de classification de 17,5% [mane07]. Et ce qui concerne les accouchements prématurés nous avons obtenu une erreur de classification de 5% avec l'USCM (tableau 5.8) et de 10% avec le CNNM (tableau 5.9). Le groupe de Garfield avait obtenu une erreur de classification de 18,5% [mane07].

Maner et al. en 2007 ont conclu que l'accouchement à terme peut être détecté 24 heures à l'avance, et que l'accouchement prématuré peut être détecté 4 jours avant la date d'accouchement. Mais d'après les résultats que nous avons obtenus dans ce présent travail, nous pouvons conclure que l'accouchement à terme peut être détecté 7 semaines à l'avance (tableau 5.8), et que l'accouchement prématuré peut être détecté 3 semaines avant la date d'accouchement (tableau 5.8).

Afin d'utiliser ces méthodes de classification supervisées, nous avons besoin d'un grand nombre de signaux pour la phase d'apprentissage. L'utilisation du modèle AR nous permet de résoudre ce problème du nombre de signaux nécessaires pour l'apprentissage. En effet, le modèle AR peut produire un grand nombre de signaux simulés d'EMG à partir d'un faible nombre de signaux d'EMG réels.

Si on compare la méthode Wavnet (supervisée) et la méthode USCM (non-supervisée) on trouve que la méthode Wavnet donne des résultats plus précis (moins d'erreur de classification). Cette méthode nécessite une grande taille de base de données pour l'apprentissage, et un temps de calcul élevé. Par contre la méthode USCM donne des résultats de classification acceptable. Elle n'a pas besoin d'apprentissage, ni des connaissances a priori à propos des signaux de base de donnée. De plus le temps de calcul est négligeable devant le temps de calcul de Wavnet.

Dans une optique d'applications cliniques, nous avons utilisé deux approches pour la séparation des signaux d'EMG. Dans la première approche, nous avons utilisé des contractions ayant le même terme à l'enregistrement (SAR) mais un terme à l'accouchement (SAA) différent (petite différence et grande différence). L'erreur de la classification est grande quand la différence entre SAA est petite. Elle diminue quand la différence entre les SAA augmente. De cette façon, nous pouvons conclure qu'il est possible de distinguer les contractions enregistrées au même terme de grossesse mais avec des termes d'accouchement différents.

C'est-à-dire, le signal d'EMG utérin peut être utilisé pour différencier entre les femmes, à un terme d'enregistrement donné, celles qui vont accoucher à terme de celles qui vont

accoucher prématurément. La deuxième approche est de classier les événements acquis avec différents termes à l'enregistrement et pour des femmes ayant le même terme d'accouchement. La séparation est possible, d'autant plus que la différence entre les termes à l'enregistrement est grande.

D'un point de vue clinique, il est donc possible de conclure :

- Qu'il est possible de distinguer, à un terme d'enregistrement donné les contractions qui amènent à des termes d'accouchement différents (1^{ère} classification). Il est donc possible d'espérer prédire la survenue d'un accouchement prématuré pour un terme d'enregistrement donné.
- Le délai entre la date d'enregistrement et la date d'accouchement a une influence sur les caractéristiques des contractions (deuxième classification), pour un terme d'accouchement donné. Il serait donc possible éventuellement de prédire la durée entre l'enregistrement et l'accouchement, et donc ainsi de prédire l'urgence de la menace.
- Les caractéristiques des contractions évoluent au cours de la grossesse, pour une même famille d'accouchement AP ou AT (deuxième étude avec accouchement AP). il est donc urgent d'augmenter la base de donnée de signaux pour fiabiliser les algorithmes de classification pour tous les termes d'enregistrement d'intérêt (de 20 à 37 SA).

Conclusion générale

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'autres parts à une définition claire du problème. Tout d'abord nous avons expliqué les MAP (Menaces d'accouchement prématuré), ses causes, et les méthodes utilisées en gynécologie-obstétrique pour détecter et prévenir les accouchements prématurés (Mesure de marqueurs biochimiques, signes cliniques, détection du rythme circadien de l'activité utérine...).

Le signal EMG semble être un vecteur potentiel d'indication du risque d'accouchement prématuré. Le signal EMG a ensuite été décrit dans ce chapitre. Il est apparu clairement que l'EMG utérin est un signal non stationnaire dont les composantes semblent riches en renseignements quant à la nature efficace de la contraction correspondante. Les caractéristiques temporelles et fréquentielles du signal ont été expliquées. A la fin de ce chapitre nous avons introduit le but de cette thèse qui peut être résumé comme une classification entre contractions à termes ou prématurés. En utilisant des méthodes de classification supervisées et non supervisées qui n'avaient pas été testées jusqu'alors pour la détection des MAP.

La classification doit être appliquée sur des représentations spécifiques du signal original. La décomposition en ondelettes a été utilisée pour extraire des paramètres pertinents. Après avoir fait la décomposition en coefficients de détails, les variances sur les échelles sont utilisées comme paramètres pour la classification.

La modélisation par modèle AR a été utilisée afin de générer une grande base de données à partir de signaux EMG utérin. Pour faciliter l'apprentissage des méthodes supervisées.

La modélisation spectrale reposant sur l'estimation par le périodogramme de Welch a été retenue pour sa capacité à produire une estimation spectrale dont il est possible de réduire l'influence d'un bruit blanc additif. Les déciles ont été utilisés comme paramètres caractéristiques de la DSP.

Le contenu fréquentiel des contractions change d'une femme à une autre et durant la grossesse. Pour cela, nous avons proposé une méthode de classification non supervisée originale basée sur la décomposition en ondelette et le test statistique et Fisher. La décomposition par ondelette est utilisée pour l'extraction des paramètres de chaque contraction. Une méthode de classification non supervisée basée sur le test de Fisher est ensuite utilisée pour classifier les événements. Une projection par Analyse en Composantes Principales permet de visualiser les groupes résultants de la méthode de classification.

Nous avons ensuite développé une nouvelle technique de classification basée sur la construction de réseaux d'ondelettes fondés sur la transformée discrète. Les réseaux obtenus sont plus volumineux, pour une même EQMP (l'erreur quadratique moyenne de performance), que ceux, fondés sur la transformée continue, qui utilisent les techniques d'optimisation non linéaire. De plus, pour des modèles de trois entrées ou plus, le nombre d'éléments dans la bibliothèque devient très grand. En revanche, cette approche présente l'avantage d'utiliser les spécificités de cette famille de fonctions pour une construction de réseaux à moindre coût (temps de calcul très inférieur à celui d'un apprentissage utilisant

une technique d'optimisation non linéaire). Cette approche est donc conseillée dans le cas où l'on désire construire un modèle rapidement sans avoir recours à plusieurs apprentissages.

Finalement, les méthodologies présentées dans ce mémoire ont été validées sur le problème de classification des signaux EMG utérin. Elles fournissent des groupes des signaux classifiés selon leur contenu fréquentiel.

En effet, afin d'employer ces méthodes de classification, nous avons besoin d'un grand nombre de signaux pour la phase d'apprentissage. L'utilisation du modèle AR nous permet de résoudre ce problème du nombre de signaux nécessaires pour l'apprentissage. En effet, le modèle AR peut produire un grand nombre de signaux simulés d'EMG à partir d'un faible nombre de signaux d'EMG réels.

D'un point de vue clinique, il est donc possible de conclure :

- Qu'il est possible de distinguer, à un terme d'enregistrement donné les contractions qui amènent à des termes d'accouchement différents. Il est donc possible d'espérer prédire la survenue d'un accouchement prématuré pour un terme d'enregistrement donné.
- Le délai entre la date d'enregistrement et la date d'accouchement a une influence sur les caractéristiques des contractions, pour un terme d'accouchement donné. Il serait donc possible éventuellement de prédire la durée entre l'enregistrement et l'accouchement, et donc ainsi de prédire l'urgence de la menace.

Les caractéristiques des contractions évoluent au cours de la grossesse, pour une même famille d'accouchement AP ou AT (deuxième étude avec accouchement AP). Il est donc urgent d'augmenter la base de donnée de signaux pour fiabiliser les algorithmes de classification pour tous les termes d'enregistrement d'intérêt (de 20 à 37 SA).

A la suite de ce travail, plusieurs directions de recherche, qui n'ont pas été encore abordées ou approfondies, nous semblent intéressantes à poursuivre. Il est possible tout d'abord d'améliorer la méthode de classification statistique non supervisée (USCM) par utilisation de plus de signaux réels, afin de valider la classification obtenue, d'un point de vue clinique. Notons que la méthode USCM peut être appliquée à d'autres signaux que l'EMG utérin, en particulier pour résoudre tout problème de classification sans des connaissances a priori à propos des signaux utilisées.

D'autre part il faut travailler dans les réseaux d'ondelettes sur le choix de l'ondelette pertinente et surtout le nombre d'ondelettes (bibliothèque). Pour cela, nous proposons de faire une étude approfondie sur l'utilisation des réseaux d'ondelettes du point de vue classification et d'estimer l'équation de la surface qui sépare entre les classes. Il sera aussi intéressant de travailler sur la minimisation du temps de calcul.

D'un point de vue clinique, la plus grande lacune demeure le faible nombre de signaux disponibles. Il devient de plus en plus urgent de construire une réelle base de données de signaux qui pourrait être utilisée par toutes les équipes. Cette base de donnée ne pourra se construire que par un travail multicentrique impliquant tous les centres cliniques concernés par le diagnostic des MAP.

Référence

- [abry97] Abry P., Ondelettes et turbulences, Diderot Editeur, Arts et sciences, Paris, 1997.
- [acog88] ACOG, Ultrasound in pregnancy, American College of Obstetricians and Gynecologists, Technical bulletin, n°116, mai 1988.
- [akay96] Akay M. Detection and estimation methods for biomedical signals. Academic press, San Diego, California, 1996.
- [audr84] Audre P., Technique de maturation cervicale au premier et deuxième trimestre de la gestation, conférence sur la maturation et la dilatation du col, Lyon, France, 1992.
- [baro97] Baron R. "Contribution à l'Étude des Réseaux d'Ondelettes." Thèse de Doctorat de l'École Normale Supérieure de Lyon, 1997.
- [barr93] Barron A. R. "Universal Approximation Bounds for Superpositions of a Sigmoidal Function." IEEE Transactions on Information Theory IT-39, pp. 930-945, 1993.
- [bart48] Bartlett M. Smoothing periodograms from time series with continuous spectra. Nature, 1948, vol. 161, pp. 686-687.
- [beno99] Benoît S., Accouchement prématuré: quand plane la menace..., Gyn. Obs., 399 : 4-7, 1999.
- [bish95] Bishop C. 'Neural Networks for Pattern Recognition' Oxford University Press ISBN:0198538642, 1995.
- [blon97] Blondel B., Bréart G., Mazaubrun C., Badeyan G., Wcislo M., Lordier A., Matet N., La situation périnatale en France. Évolution entre 1981 et 1995. J Gynecol Obstet Biol Reprod., 26 : 770-80, 1997.
- [blon00] Blondel B., Norton J., du Mazaubrun C., Bréart G., Enquête nationale périnatale 1998. Rapport Inserm, 2000.
- [blon01] Blondel B., Norton J., du Mazaubrun C., Bréart G., pour la coordination nationale des enquêtes nationales périnatales. Évolution des principaux indicateurs de santé périnatale en France métropolitaine entre 1995 et 1998. Résultats des enquêtes nationales périnatales, J Gynecol Obstet Biol Reprod, 30 : 552-64, 2001.
- [bode31] Bode O., Das elektrohysterogram, Arch. Gynaekol., 153: 123, 1931.
- [brea95] Bréart G., Blondel B., Tuppin P., Grandjean H., Kaminski M., Did preterm deliveries continue to decrease in France in the 1980s?, Paediatr Perinat Epidemiol, 9: 296-306, 1995.
- [bren90] Brennecke D., Rice G. et Scott K., Elevation of peripheral plasma Phospholipase A2 in association with human preterm labour, Proceedings of the 29th National Scientific Conference of the Australian Society for Medical Research, 1990.
- [bruc85] Bruckstein A. M. The resolution of overlapping echo's. IEEE Trans. Acoustic speech. signal. process., vol. ASS SP -33, 1985, p. 1357.

- [bui96] Buhimschi C., et Garfield R.E. Uterine contractility as assessed by the abdominal surface recording of EMG activity in rats during pregnancy. *Am. J. Obstet. Gynecol.* 174: 744-53;1996.
- [bui97] Buhimschi C., Boyle M.B., et Garfield R.E. Electrical activity of the human uterus during pregnancy as recorded from the abdominal surface. *Obstet. Gynecol.* 90:102-11;1997.
- [bui98] Buhimschi C., Boyle M.B., Saade G.R., et Garfield R.E. Uterine activity during pregnancy and labor assessed by simultaneous recordings from the myometrium and abdominal surface in the rat. *Am. J. Obstet. Gynecol.* 178:811-22;1998.
- [cald80] Calder A., Pharmacological management of the unripe cervix in the human, in 'Dilatation of the uterine cervix', Naftolin F. et Stubblefield P. Eds, Raven Press, New York, 317-333, 1980.
- [camp87] Campbell E., Linton E. et Wolfe C., Plasma corticotrophin releasing hormone concentration during pregnancy and parturition, *J. Clin. Endocrinol. Meta.*, 34:1054, 1987.
- [cann95] Cannon M. et Slotine J.-J. E. Space-Frequency Localized Basis Function Networks for Nonlinear System Estimation and Control, *Neurocomputing* 9 (3) (1995) 293-342.
- [carm97] Carmichael S., Abrams B., Selvin S. The association of pattern of maternal weight gain with length of gestation and risk of spontaneous preterm delivery. *Paediatr. Perinat. Epidemiol.* 11;4 : 392-406; 1997.
- [chal93] Chalubinski K., Deutinger J., et Bernaschek G., Vaginosonography for recording of cyclereleted myometrial contraction. *Fert. And Ster.*, 59 :1: 225-228, 1993.
- [chen89] Chen S., Billings S.A. et Luo W. "Orthogonal Least Squares Methods and Their Application to Non-linear System Identification." *Int. Journal of Control*, Vol. 50, No. 5, pp. 1873-1896, 1989.
- [chke07] Chkeir A. Classification des signaux électromyographies utérins par réseaux d'ondelettes. Master de recherche de l'Université de Technologie de Compiègne. 2007.
- [cich93] Cichocki A., Unbehauen R. *Neural Networks for Optimization and Signal Processing*. Jhon Wiley & Sons Ltd & B. G. Teubner, Stuttgart, 1993.
- [coat83] Coatrieux J.L. Interference electromyogram processing. Part 2: Experimental and simulated EMG AR modeling. *Electromyogr. Clin. Neurophysiol.*, vol. 23, 1983, pp. 481-490.
- [copp90] Copper R. Goldenberg R., Dubard M., Hauth J. et Cutter G., Warning symptoms uterine contraction, and cervical findings in women et risk of preterm delivery, *Am. J. Obstet. Gynecol.*, 162:748, 1990.
- [crea84] Creazy R., et Resnik R., *Maternal fetal Medicine : Principles and Practice*, Saunders WB. Ed., Philadelphia, 420, 1984.
- [darn87] Darne J., Mcgarrigle H. et Lachelin G., Increased saliva oestriol to progesterone ratio before preterm delivery : A possible predictor for preterm labor ?, *Br. Med. J.*, 294:270, 1987.
- [daub92] Daubechies I. "Ten Lectures on Wavelets." *CBMS-NSF regional series in applied mathematics*, SIAM, Philadelphia, 1992.
- [deve93] Devedeux D., Marque C., Mansour S. et al., Uterine EMG : a critical review, *AM. J. Obstet. Gynecol.*, 169 :1636-1653, 1993.

- [deve95a] 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.
- [deve95b] Devedeux d., Duchene J., Germain G. et al., Evolution Temporelle des principales composantes spectrales de l'électromyogramme utérin : étude expérimentale. *Innov. Techn. Biol. Med.*, 16 : 4: 496-480, 1995.
- [diab04a] Diab M.O., Marque C., Khalil M. An unsupervised classification Method of uterine Electromyography Signals using wavelet decomposition. Proceedings of the 26th Annual International Conference of the IEEE EMBS, San Francisco, CA, USA, September 1-5. 192-195. 2004.
- [diab04b] Diab M.O., Marque C., Khalil M. Uterine Electromyography Signals: Classification and Interpretation. Proceedings of the 2nd Cairo International Biomedical Engineering Conference (CIBEC, IEEE EMBS), Cairo, Egypt, December 27-28. 1-4. 2004.
- [diab05] Diab, M.O., Marque C., Khalil M. Unsupervised Classification in Uterine Electromyography Signal: Toward The Detection of Preterm Birth. Proceedings of the 27th Annual International Conference of the IEEE EMBS, Shanghai, China, September 1-4. 2232-2235. 2005.
- [diab06a] Diab M. O., Marque C., Khalil M. Une approche de classification des contractions utérines basées sur la théorie des ondelettes et la statistique. *Leb. Sc. J. CNRSL press.* 6; 1: 91-103; 2006.
- [diab06b] Diab M. O., Marque C., Khalil M. Modelling And Classification For Uterine EMG Signals Using Autoregressive Model. *Intellig. Eng. Sys. A. N. N.*, ASME press. 8;1: 317-23, 2006.
- [dill46] Dill L., et Maiden R., The electrical potentials of the human uterus in labor. *Am. J. Obstet. Gynecol.*, 52:735-45, 1946.
- [dree 98] Direction de la Recherche des Etudes de l'Evaluation et des Statiques. Chiffres repères. Ministère de l'emploi et de la solidarité, 1998.
- [drey02] Dreyfus G., Martinez M., Gordon M. Réseaux de neurones. Eyrolles. 2002.
- [dutch93] Dûchene J., Goubel F. Surface Electromyogram during voluntary contraction: processing tools and relation to physiological events, *Crit. Rev. in Biomed. Eng.*, 21: 4: 313-397, 1993.
- [dutch95] Dûchene J., Devedeux D., Mansour s. et al. Analysing uterine EMG : Tracking instanteneous burst frequency, *IEEE Eng. In Medicine and Biology Magazine*, 125-131, 1995.
- [dupu99] Dupuis O., Thoulon J. et Mellier G., Apport de l'échographie du sol dans la prédiction de l'accouchement prématuré : revue de la littérature, *Contracep. Fert. Es.*, 27 :2 :147-154,1999.
- [falo02] 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 Troyes, France, Décembre 2002.
- [fish73] Fisher W., *Kardiotokographie Lehrbuch und Atlas*, Thieme-Verlag, Stuttgart 1973.
- [flan93] Flandrin P. temps fréquence, édition Hermès, Paris, 1993.
- [floy61] Floyd W., Cervical dilation in the mid trimester of pregnancy. *Obstet. Gynecol.*, 18:30, 1961.

- [fuch73] Fuchs A., Hormonal control of myometrial function during pregnancy and parturition, *Acta End.*, 221: 1-69, 1973.
- [garf80] Garfield R., Kannan M. et daniel E., Gap junction formation in myometrial: control by estrogens, progesterone and prostaglandins, *Am. J. Physiol.*, 238-281, 1980.
- [garf84] Garfiel R., Myometrial ultrastructure and uterine contractility. In: *Uterine Contractility*, edited by Bottari S., Thomas J.P., Vokaer A., Vokaer R. New York: Masson, 81-109, 1984.
- [garf98] Garfield RE., Buhimschi C., Buhimschi I., Shi L., Schi SQ. et al. Control and assessment oh the uterus and cervix during pregnancy and labour. *Hum Reprod Update.* 4:673-95;1998.
- [garf01] Garfield R.E., Maul H., Shi L., Maner W., Fittkow C., Olsen G., et al. Methods and devices for the management of term and and preterm labor. *Ann. N. Y. Acad. Sci.* 943:203-24; 2001.
- [geor00] Georgiou T.T Signal estimation via selective harmonic amplification: MUSIC, Redux . *IEEE Trans. On signal Processing*, vol.48, March 2000, pp. 780-790.
- [germ93] Germain A., Valenzuela G., Ivankovic M., Ducsay C., Gabella C. et Seron Ferre M., Relationship of circadian rythms of uterine activity with term and preterm delivery, *Am. J. Obstet. Gynecol.*, 168:1271, 1993.
- [goff01] Goffinet F., Maillard F., Fulla Y., Cabrol D., Biochemical markers (without markers of infection) on the risk of preterm delivery. Implications for clinical practice, *Eur J Obstet Gynecol Reprod Biol*, 94: 59-68, 2001.
- [gova03] Govaert G, "Analyse des données", Hermes Lavoisier, Paris, 2003.
- [grau89] Graupe D., Kohn K.H., Basseas S. Control of electrically-stimulated walking of paraplegics via above and below lesion EMG signature definition. *IEEE Trans. Auto .Control.*, vol.34, 1989, p 130.
- [guin94] Guinn D., Wigton T., Owen J., Socol M. et Frederiksen M., Prediction of preterm birth in nulliparous patients, *Am. J. Obstet. Gynecol.*, 171:1111, 1994.
- [haas96] Haas JS., Berman S., Goldberg AB., Lee LWK., Cook EF., Prenatal hospitalization and compliance with guidelines for prenatal care, *Am J Public Health*, 86: 815-9, 1996.
- [hall89] Hallman M., Bry K. et Pitkanen O., Ceramide lactoside in amniotic fluid: high concentration in chorioamnionitis and in preterm labor, *Am. J. Obstet. Gynecol.*, 161: 313, 1989.
- [heam01] Heaman M.I., Sprague A.E., Stewart P.J., Reducing the preterm birth rate: a population health strategy, *JOGNN*, 30: 20-9, 2001.
- [hebb49] Hebb D. O. *The Organization of Behavior*. New York: John Wiley, 1949.
- [heff88] Heffner G., Jaros J. The electromyogram (EMG) as a control signal for functional neuromuscular stimulation. II. Practical demonstration of the EMG signature discrimination system. *IEEE Trans. Bio. Eng.*, vol.34, 1988, p 230.
- [hogr94] Hogrel J.Y. *Analyse du signal electromyographique pour la caractérisation d'individus et la discrimination de population*. Thèse de doctorat de l'Université de Technologie de Compiègne. 1994.
- [hon58] Hon E.H.G. et Davis C.D., Cutaneous and utérine electrical potentials in labor, *Obstet. Gynecol.*, 12: 47, 1958.

- [horn94] Hornik K., Stinchcombe M., White H. et Auer P. "Degree of Approximation Results for Feedforward Networks Approximating Unknown Mappings and Their Derivatives." *Neural Computation*, Vol. 6, No. 6, pp.1262-1275, 1994
- [horn89] Hornik K., Stinchcombe M. et White H. "Multilayer feedforward networks are universal approximators." *Neural Networks* 2, pp. 359-366, 1989.
- [hues98] Hueston WJ., Preterm contractions in community settings: II. Predicting preterm birth in women with preterm contractions, *Obstet Gynecol*, 92: 43-6, 1998.
- [inba84] Inbar G.F., Allin J., Paiss O. et al., Monitoring surface EMG spectral changes by the zero crossing rate, *Med. & Biol. Eng. & Comput.*, 31:597, 1984.
- [judi94] Juditsky A., Zhang Q., Delyon B., Glorennec P. Y. et Benveniste A. "Wavelets in Identification: wavelets, splines, neurons, fuzzies: how good for identification?" Rapport INRIA No. 2315, Septembre 1994.
- [kao77] Kao C.Y., Electrophysiological properties of the uterine smooth muscle, in *Biology of the uterus*, Wynn. R.M. Ed., New York, Plenum Press, 423-496, 1997.
- [katz86] Katz M., Newman R. et Gill P., Assessment of uterine activity in ambulatory patients at high risk of preterm labor and delivery, *Am. J. Obstet. Gynecol.*, 154:44-7, 1986.
- [katz90] Katz M., Goodyear K. et Creasy R., Early signs and symptoms of preterm labor, *Am. J. Obstet. Gynecol.*, 162:1150, 1990.
- [kay88] Kay S. *Modern Spectral Estimation Theory & Application*. Englewood Cliffs, Prentice-Hall, NJ, USA, 1988.
- [kerj99] Kerjosse R., Bilan démographique 1998: natalité de nouveau en hausse. *INSEE Première*, 633,1999.
- [khal97] Khalil M., Duchêne J. et Marque C., Reassignment method in time-scale representation. Application to uterine electromyography. *World Congress on Medical Physics and Biomedical Engineering*, Nice, France, 557, 1997.
- [khal99] 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. 1999.
- [king88] King JF., Grant A., Keirse MJ., Chalmers I., Beta-mimetics in preterm labor: an overview of the randomized controlled trials, *BJOG*, 95: 211-22, 1988.
- [kloc75] Klöck F. et Chantraine H., Möglichkeiten der Behandlung einer drohenden Frühgeburt, *Gynäkologe*, 9:206, 1975.
- [koep75] Koepke E., Herrmann C. et Seidenschur G., Ambulante Tokographie, *Zbl Gynäkologe*, 9:206, 1975.
- [koro99] Korosec D., Zazula D., Merletti R. Influence of the linear model structure and the electrode placement to the estimation of the mean frequency in surface EMG signals. *Proc. of the fourth general SENIAM workshop's*, Hertogenbosh, Netherlands, September 1999, pp. 31- 37.
- [krim95] Krim H., Pesquet J.C. Multiresolution analysis of a class of non stationary processes. *IEEE Trans. Information Theory*. 41 : 4, 1011-1020. 1995.

- [kuga95] Kugarajah T. et Zhang Q. "Multidimensional Wavelet Frames." IEEE Trans. on Neural Networks, Vol. 6, No. 6, pp. 1552-1556, November 1995.
- [kuma61] Kumar D. et Barnes C., Studies in human myometrium during pregnancy. II resting membrane potential and comparative electrolyte levels. Am. J. Obstet. Gynecol., 82-736-741, 1961.
- [kunt91] Kunt M., Bellanger M., De Coulon F., Gueguen C., Hasler M., Moreau N., Vetterli M. Techniques modernes de traitement numérique des signaux. Presses polytechniques et universitaires romandes, Lausanne, 1991.
- [kuri61] Kuriyama H. et Cspa A.I., A study of the parturient uterus with the microelectrode technique, Endocrinology, 68 :1010-25, 1961.
- [lang62] Langreder W., Die externe Tokometrie, VEB Georg Thieme, Leipzig, 1962.
- [lema96] Leman H. Paramétrage du signal EMG utérin pour la détection précoce des menaces d'accouchement prématuré. DEA de l'Université de Technologie de Compiègne. 1996.
- [lema97] Leman H., Marque C., Optimum denoising method for the uterine EHG, XIX th IEEE-EMBS Conf., Chicago, 28 Oct.-02 Nov. 1997.
- [lema98] Leman H., Marque C. Ridge extraction from the scalogram of the uterine electromyogram, IEEE TFTS, 1998.
- [lema99] Leman H. Apport des ondelettes dans le traitement de l'électromyogramme utérin abdominal: caractérisation des contractions pendant la grossesse pour la détection des menaces d'accouchement prématuré. Thèse de doctorat de l'Université de Technologie de Compiègne. 1999.
- [leve86] Leveno K., Cox K., Senyei A. et Dische M., Cervical dilatation and prematurity revisited, Obstet. Gynecol., 68:434, 1986.
- [li95] Li C., Zheng C., Tai C. Detection of ECG characteristic points using wavelet transform. IEEE transaction, BME. 42: no. 1, 21- 29. 1995.
- [ligg79] Liggins G., Initiation of parturition, Br. Med. Bull., 35:145-150, 1979.
- [lind80] Linde Y., Buzo A., Gray R.M. "An algorithm for vector quantization (QV) design". IEEE Transactions on Communication. Com-28 :1 ; 84-95, January 1980.
- [lope84] Lopes P., Germain G., Breart G., Reinato S., Le Houezec R. et Sureau C., Electromyographical study of uterine activity in the human during labor induced by prostaglandin F2 alpha, Gynecol. Obstet. Invest., 17: 96, 1984.
- [mane03] Maner W.L., Garfield R.E, Maul H., Olson G., et Saade G. Predicting term and preterm delivery with transabdominal uterine electromyography. Obstet. Gynecol. 101; 6:1254-60; 2003.
- [mane07] Maner W.L., et Garfield R.E. Identification of human term and preterm labor using artificial neural networks on uterine electromyography data. Ann. Biomed. Eng. 35; 3:465-73; 2007.
- [mans91] Mansour S., Duchêne J., Germain G. et Marque C., Utérine EHG : Experimental and mathematical détermination of the relationship between internal and external recordings, in Proc. 13th Annu. IEEE EMBS Conf., Orlando, USA, 1991.

- [mans93] Mansour S., Etude de l'électromyogramme utérin: caractérisation, propagation, modélisation des transferts, Thèse de l'Université de Technologie de Compiègne, Spécialité GBM, 1993.
- [marp87] Marple S. L. Digital spectral analysis with applications. Prentice-Hall, Englewood Cliffs, NJ, USA, 1987.
- [marq86] Marque C., Dûchêne J., Leclercq S., Panczer G., Chaumont J. Uterine EHG processing for obstetrical monitoring. IEEE Trans. BME. 33 : 1182-1187. 1986.
- [marq87] Marque C., Analyse de la dynamique des contractions utérines par électromyographie abdominale, Thèse de l'Université de Compiègne, spécialité GBM, 1987.
- [marq89] Marque C. et Duchêne J., Human abdominal EHG processing for uterine contraction monitoring, in Applied biosensors, Wise DL., Ed., Stoneham, Butterworth, 187-226, 1989.
- [marq95] Marque C, Gondry J, Rossi J, Baaklini N et Duchêne J, Surveillance des grossesses à risque par électromyographie utérine, RBM,17 : 25, 1995.
- [mars84] Marshall J., Comparison of circular and longitudinal activity in the pregnant myometrium, in uterine contractility, Bottari S. et Thomas J. Ed. Masson, New York, 127-135, 1984.
- [mcle93] Mclean M., Walters W. et Smith R., Prediction and Early Diagnosis of Preterm Labor :A critical Review, Obstetrical and Gynecological Survey, 48:209,1993.
- [merg79] Merger R., Levy J., et Melchior J., Précis d'obstétrique, Masson, Paris, 1979.
- [merl99a] 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, 1999, pp. 65-78.
- [merl99b] Merletti R., Hermens H. J., Freriks B., Hagg G., Disselhorst-Klug C. Conclusions of the SENIAM signal processing task. SENIAM- Deliverable 7, State of the art on signal processing methods for SEMG, 1999, pp. 245 –250.
- [nage83] Nagel J. et Schaldach M., The non-traumatisante assessment of uterine activity, Non-traumatisante measurements, 2:103-129, 1983.
- [nage88] Nageotte M., Quantification of uterine activity preceding preterm, term and postterm labor. Am. J. Obstet. Gynecol., 158:1254, 1988.
- [nare90] Narendra K. S. et Parthasarathy K. "Identification and Control Of Dynamical Systems Using Neural Networks." IEEE Trans. on Neural Networks Vol.1, pp. 4-27, 1990.
- [newm87] Newman R., Antepartum Ambulatory tocodynamometry : The significance of Low-Amplitude, high-Frequency contraction, Obstet. Gynecol., 70:701, 1987.
- [olse95] Olsen P., Läärä E., Rantakallio P., Järvelin MR., Sarpola A., Hartikainen AL., Epidemiology of preterm delivery in two birth cohorts with an interval of 20 years. Am J Epidemiol., 142: 1184-93, 1995.
- [orfa90] Orfanidis S. J. Optimum Signal Processing an Introduction, McGraw-Hill international editions, Singapore, 1990.

- [must03] Mustafa O. Caractérisation des signaux EMG de longue durée par les réseaux d'ondelettes. DEA de l'Université Libanaise en collaboration avec l'Université de Reims, Université de Rennes et l'Ecole Polytechnique Fédérale de Lausanne. 2003.
- [ouss98] Oussar Y. Réseaux d'ondelettes et réseaux de neurones pour la modélisation statique et dynamique de processus. Thèse de doctorat de l'Université Pierre et Marie Curie. 1998.
- [papi86] Papiernik E., Bouyer J. et Collin D., Precocious cervical ripening and preterm labor, *Obstet. Gynecol.*, 67:238, 1986.
- [pati93] Pati Y. C. et Krishnaprasad P. S. "Analysis and Synthesis of Feedforward Neural Networks Using Discrete Affine Wavelet Transformations." *IEEE Trans. on Neural Networks* Vol. 4, No. 1, pp. 73-85, 1993.
- [park91] Parks J., et Sandberg I.W. 'Universal Approximation using Radial-Basis Function Network' *Neural Computation* V3:n2,PP246-257, 1991.
- [pari61] Parikh M. et Metha A., Internal cervical os during the second half of pregnancy, *J. Obstet. Gynecol. Br. Commonw.*, 68:818, 1961.
- [parl84] Parlatti E., Dell'acqua S., Cinque B., Travaglini R., Navaretta V., Maniccia E., et Bompiani A., Endocrine factors involved in the mechanism of labor in humans, in 'uterine contractility', Bottari S., Thomas J.P. et Vokaer R. Eds, masson Publishing, USA, Inc, 345-352, 1984.
- [peac97] Peaceman A., Fetal fibronectine as a predictor of preterm birth in patients with symptoms: A multicenter trial, *Am. J. Obstet. Gynecol.*, 177: 13-18, 1997.
- [plan84] Planes J., Favretto R., Gransjean H. et Morucci J., External recording and processing of fast electrical activity of the uterus in human parturition, *Med. & Biol. Eng. & Comp.*, 22:585-591, 1984.
- [pogg89] Poggio T., et Girosi F. 'Theory of Networks for Approximation and Learning' *A. I. Memo*, 1140,1989.
- [proa96] Proakis J. G., Manolakis D. G., *Digital signal processing principles, Algorithms, and Applications.* Englewood Cliffs, Prentice- Hall, NJ, 1996.
- [redd79] Reddi S. S. Multiple source location – A digital approach. *IEEE Trans. Aero. Electro. Syst.*, vol. AES -15, 1979, pp. 95-105.
- [rome89] Romero R., Brody D. et Oyarzun E., Infection and labor. III. Interleukin-1: A signal for the onset of parturition, *Am. J. Obstet. Gynecol.*, 160: 1117, 1989.
- [rose62] Rosenblatt F, "Principles of neurodynamics", Spartan Press, East Lansing, MI, 1962.
- [rous 98] Roussel S., Levenberg-Marquardt optimization, <http://www.gatsby.ucl.ac.uk/~raweis/notes.html>.
- [rume86] Rumelhart D. E. et McClelland J. L. "Parallel Distributed Processing,." MIT Press, Cambridge, MA, 1986.
- [rutt72] Rüttgers H., Kubi F. et Haller G., Die antepartale foetale Herzfrequenz, I. Mitteilung. *Z. Geburtsh. Perinat.*, 176:294, 1972.
- [sali82] Saling E., Praematuritäts und Dysmaturitäts-Präventivprogramm, *Z. Geburtsh. Perinat.*, 176:70, 1982.

- [sapo90] Saporta G. "Probabilités, Analyse des données et Statistiques". Editions Technip, Paris; 1990.
- [sena99] Senat M.V., Tsatsaris V., Ville Y. et Fernandez H., Menace d'accouchement prématuré. Encycl. Méd. Chir., Elsevier, Paris, Urgences, 17 p; 1999.
- [seid87] Seidel H., Beyer H., Brauer D. Electromyographic evaluation of back muscle fatigue with repeated sustained contractions of different strengths. Eur. J. Appl. Physiol., 56, 592. 1987.
- [sesi 98] Service des Statistiques, des Etudes et des Systèmes d'Information. Comptes nationaux de la santé, Ministère de l'emploi et de la solidarité, 1998.
- [shen95] Shenhadji L., Bellanger J.J., Carraut G., Détection temps échelle d'événements paroxystiques intercritiques en électroencéphalogramme, Traitement du Signal, 1995 ; 12 ; 4 : 357-371.
- [sims82] Sims S. M., Daniel E.E. et Garfiel R., Improved electrical coupling in uterine smooth muscle in associated with increased numbers of gap junction at parturition. J. Gen. Physiol., 80:353-375, 1982.
- [stee50] Steer CM. et Hertsh GJ., Electrical activity of the human uterus in labor, Am. J. Obstet. Gynecol., 68: 867, 1950.
- [sure56] Sureau C, Etude de l'activité électrique de l'utérus au cours du travail, Gynecol. Obstét., 55: 153, 1956.
- [tert03] Tertois S. Réduction des effets des non-linéarités dans une modulation multiporteuse à l'aide de réseaux de neurones. Thèse de doctorat de l'Université de Rennes 1. 2003.
- [ther92] Therrien W. C. Discrete random signals and statistics signal processing. Englewood Cliffs, Prentice-Hall, NJ, 1992.
- [torr95] Torrèsani B. "Analyse Continue par Ondelettes." InterEditions / CNRS Editions, Paris 1995.
- [tort94] Tortora G., et Grabowski S., Principes d'anatomie et de physiologie, De Boeck Université, Québec, 1994.
- [val79] Val N., Bubuissou B. et Goubel F., Aide au diagnostic de l'accouchement par l'électromyogramme abdominal : sélection de caractères. Reconnaissance des formes et intelligence artificielle, 3 :42-8, 1979.
- [verd01] Verdenik I., Pajntar M., et Leskosek B. uterine electrical activity as predictor of preterm birth in women with preterm contractions. Eur. J. Obstet. Gynecol. Reprod. Biol. 95 : 149-153, 2001.
- [vois00] Voisine ML, 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, Toulouse, mars, 22, 2000.
- [vois01] Voisine M.L., Réalisation et validation clinique d'un prototype d'appareillage pour le suivi à domicile de l'efficacité des contractions utérines pendant la grossesse. Thèse de doctorat de l'Université de Technologie de Compiègne. 2001.
- [vois02] Voisine M.L., Marque C., Leman H., Gondry J. Analysis of uterine electromyogram for the monitoring of preterm birth. Instrumentation development and data analysis. 4th International Workshop on Biosignal Interpretation, Como, Italie, 75-78, Juin 2002.

- [wolf79] Wolfs G., et Van Leeuwen M., Electromyographic observations on the human uterus during labour, Acta. Obstet. Gynec. Scand., 90:1-61, 1979.
- [wolf70] Wolfs G. et Rottinghuis H., Electrical and mechanical activity of the human uterus during labour, Arch. Gynaekol., 208: 373, 1970.
- [wood72] Wood C., Myometrial and tubal physiology, in 'human reproductive physiology', Sheraman R. Ed., Oxford, England, Blackwell, 324-375, 1972.
- [zahn84] Zahn V., Uterine contraction during pregnancy, J. Perinat. Med., 12:107, 1984.
- [zhan92] Zhang Q. et Benveniste A. "Wavelet Networks. "IEEE Trans. on Neural Networks Vol. 3, No. 6, pp. 889-898, 1992
- [zhan93] Zhang Q. "Regression Selection and Wavelet Network Construction." Rapport interne de l'INRIA N. 709, Projet AS, Avril 1993
- [zhan95] Zhang J., Walter G. G., Miao Y. et Wayne Lee W. N. "Wavelet Neural Networks For Function Learning." IEEE Trans. on Signal Processing, Vol. 43, no. 6, pp. 1485-1497, 1995
- [zhan97] Zhang Q. "Using Wavelet Network in Nonparametric Estimation." IEEE Trans. on Neural Networks, Vol. 8, No. 2, pp. 227-236, 1997.
- [zhan00] Zhang Q. J., Gupta K. C. "Neural Networks for RF and Microwave Design". Artech House Publishers; Har/Cdr edition, 2000.

Publications relatives à ce travail

Article journal

1. **Diab M. O.**, Marque C., Khalil M. An Unsupervised Classification Method Of Uterine Electromyography Signals: Classification For Detection Of The Preterm Deliveries. J. Obstet. Gynecol. Res., accepted Octobre 2007.
2. **Diab M. O.**, Marque C., Khalil M. Classification for Uterine EMG Signals: Comparison Between AR Model And Statistical Classification Method. Int. J. Congnition and Computational, vol. 5, no. 1, 8-14, March 2007.
3. **Diab M. O.**, Marque C., Khalil M. Modeling And Classification For Uterine EMG Signals Using Autoregressive Model. Intelligent Engineering Systems through Artificial Neural Networks, ASME press. Vol. 8, No. 1: 317-323, 2006.
4. **Diab M. O.**, Marque C., Khalil M. Une approche de classification des contractions utérines basées sur la théorie des ondelettes et la statistique. Journal of National Council for Scientific Research. Vol. 6, No. 1: 91-103 2006.

Conférence International avec Comité de Lecture

5. **Diab M. O.**, Marque C., Khalil M. Modeling And Classification For Uterine EMG Signals Using Autoregressive Model. ANNIE conference, Novembre 2006.
6. **Diab, Mohamad O.**; Marque, Catherine; Khalil, Mohamad. Unsupervised Classification in Uterine Electromyography Signal: Toward The Detection of Preterm Birth. 27th Annual International Conference of the IEEE EMBS Shanghai, Chine, 4-8 September 2005, p: 230-236.
7. **Diab M. O.**, Marque C., Khalil M. An Unsupervised Classification Method of Uterine Electromyography Signals Using Wavelet Decomposition. Proceedings of the 26th Annual International Conference of the IEEE EMBS San Francisco, CA, USA, September 1-5, 2004, p: 192-195.
8. **Diab M. O.**, Marque C., Khalil M. Uterine electromyography signals: Classification and Interpretation. Proceedings of the 2nd Cairo International Biomedical Engineering Conference (CIBEC, IEEE EMBS), Cairo, Egypt, December 27-28, 2004, p: 1-4.