

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediène
Faculté d'Électronique et d'Informatique



THÈSE

Présentée pour l'obtention du **grade** de **DOCTEUR EN SCIENCES**

En : ÉLECTRONIQUE

Spécialité : Instrumentation Électronique

Par: OULD EL HADJ Saliha Epouse **OULD SLIMANE**

THEME

**Modélisation et Compression du Signal Cardiaque
en Vue d'une Détection Automatique
des Anomalies Cardiovasculaires**

Soutenue publiquement le **9 / 12 / 2017**, devant le jury composé de :

Mr R. OUSSAID	Professeur à l'U.S.T.H. B	Président
Mme M. TALHA-KEDIR	Professeure à l'U.S.T.H.B	Directrice de thèse
Mme M. BOUDRAA	Professeure à l'U.S.T.H.B	Examinatrice
Mr M. MEHENNI	Professeur à l'ENP	Examineur
Mme T. DJAMAH	Maître de Conférences / A à l'UMMTO	Examinatrice
Mr A. GOUTAS	Maître de Conférences / A à l'U. Skikda	Examineur

Remerciements

Les mots ne pourraient jamais exprimer suffisamment ma sincère reconnaissance, ma gratitude et mon profond respect envers ma directrice de thèse, Professeure M.D. Talha-Kedir chef d'équipe de Génie-Biomédical du laboratoire LINS de la Faculté d'Electronique et d'Informatique de l'USTHB. Je tiens à la remercier du fond de mon cœur pour sa permanente disponibilité, pour sa générosité scientifique, sa clairvoyance, sa maîtrise du traitement de signal et le temps qu'elle m'a dispensé pour mener à terme ma thèse. Je n'oublierai pas de souligner son côté humain et chaleureux qui m'ont toujours réconfortée et soutenue dans les moments les plus difficiles. Merci encore.

J'adresse mes très sincères remerciements au professeur OUSSAID Rachid pour l'honneur et l'intérêt qu' il a accordé à ce travail en acceptant de le juger et de présider le jury.

Je tiens également à exprimer mes profonds remerciements à l'égard des membres du jury, les professeurs Boudraa Malika , Mehenni Mohamed ainsi que les docteurs Djamah Tounsia et Goutas Ahcène, qui m'ont fait l'honneur d'examiner et évaluer mon travail. Merci encore.

Je dédie ma thèse à mes très chers parents, à mon mari Salah, à mes enfants Yasmine, Mehdi et Abderrahmane, à mon beau-fils Abdennour ainsi qu'à mon adorable petit-fils Aghilès Abdelmalek. Merci pour leur soutien moral tout au long de la réalisation de ma thèse.

ملخص

هدفنا هو تصميم نظام المساعدات التشخيصية، من خلال النمذجة الرياضية من نبضات القلب وتصنيفها. لتحقيق هذا الهدف، قمنا بتطوير نهجين لاستخراج من إشارات تخطيط القلب الحقيقي، والمعايير الأكثر أهمية وكبيرة التي يمكن أن تصنف أفضل تخطيط القلب إلى فئتين: عادي وغير طبيعي. يستخدم النهج الأول الشبكات العصبية RBF (شعاعي أساس وظيفة) جنباً إلى جنب مع خوارزمية GOFR (المعمم متعامد إلى الأمام الانحدار)، لاستخراج المعلومات الزمنية التي تميز كل نبضة قلب. في حين أن الطريقة الثانية هي طريقة تردد على أساس الموجات تحويل DWT (متفصلة تحويل الموجات) حيث تم احتساب الفرق من الطاقة من معاملات التفاصيل من المستويات الخمسة الأولى من التحلل من كل إشارة ECG لتمثيل التوزيع والحفاظ على أكثر ملاءمة. في كلا النهجين، تم الجمع بين أساليب الحد من أكب أو إيكاً المعلومات مع نفس دعم ناقلات آلة (سقم) المصنف. وأظهرت النتائج التي تم الحصول عليها تناسب جيد من المعلومات ذات الصلة لكل نهج وكثفت عن الجمع بين أكثر كفاءة من الخوارزميات، والتي حققت معدلات تصنيف تصل إلى 100%.

Résumé

Notre but est de concevoir un système d'aide au diagnostic, moyennant la modélisation mathématique des battements cardiaques et leur classification. Pour atteindre cet objectif, nous avons développé deux approches pour extraire à partir des signaux ECG réels, les paramètres les plus pertinents et les plus significatifs qui peuvent classer les ECG en deux classes : Normal et Anormal. La première approche utilise Les réseaux de neurones RBF (Radial Basis Function) combinés à l'algorithme GOFR (Generalized Orthogonal Forward Regression), pour l'extraction de paramètres temporels caractérisant chaque battement cardiaque. Tandis que la seconde approche est une méthode fréquentielle basée sur la transformée en ondelette DWT (Discrete Wavelet Transform) où la variance de l'énergie des coefficients de détails des cinq premiers niveaux de décomposition de chaque signal ECG, a été calculée pour représenter leur distribution et garder les plus pertinents. Dans les deux approches, les méthodes de réduction de paramètres ACP ou ICA ont été combinées au même classifieur SVM (Support Vector Machine). Les résultats obtenus ont montré un bon ajustement des paramètres pertinents pour chaque approche et ont révélé la combinaison d'algorithmes la plus efficace, qui a permis d'atteindre des taux de classification allant jusqu'à 100%.

Abstract

Our goal is to design a diagnostic aid system, through mathematical modeling of heart beats and their classification. To achieve this goal, we have developed two approaches to extract from real ECG signals, the most relevant and significant parameters that can best classify ECG into two classes: Normal and Abnormal. The first approach uses Radial Basis Function (RBF) neural networks combined with the Generalized Orthogonal Forward Regression (GOFR) algorithm, for the extraction of temporal parameters characterizing each heartbeat. While the second approach is a frequency method based on the wavelet transform DWT (Discrete Wavelet Transform) where the energy variance of detail coefficients for the first five decomposition levels of each ECG signal, has been calculated to represent their distribution, and keep the most relevant ones. In both approaches, methods for reducing parameters ACP or ICA are combined with the same Support Vector Machine (SVM) classifier. The results obtained showed a good fit of the relevant parameters for each approach and revealed the most efficient combination of algorithms, which achieved classification rates of up to 100%.

TABLE DES MATIERES

Introduction générale	1
Etat de l'art	6
Chapitre 1 Le système cardiovasculaire	
Introduction	8
1.1 Le cœur	8
1.2 Fonctionnement électrique	9
1.3 Principe de l'électrocardiogramme	10
1.3.1 Dérivations des membres	11
1.3.2 Dérivations précordiales	13
1.4 L'ECG de Holter	14
1.5 Analyse de l'ECG	16
1.5.1 Ondes de l'ECG	16
1.5.1.1 l'onde P	16
1.5.1.2 le complexe QRS	16
1.5.1.3 l'onde T	17
1.5.1.4 l'onde U	17
1.5.2 Intervalles et segments de l'ECG	17
1.5.2.1 Intervalle PR	17
1.5.2.2 Segment PR	18
1.5.2.3 Segment ST	18
1.5.2.4 Intervalle QT	18
1.5.2.5 Intervalle PP	18
1.5.2.6 Intervalle RR	18
1.6 Contenu spectral de l'ECG	19
1.7 Les arythmies du cœur	19
Conclusion	21
Chapitre 2 Modélisation du signal cardiaque	
Introduction	22
2.1 Modélisation par réseaux à fonctions de base radiales (RBF)	23
2.1.1 Choix des RBF	23
2.1.2 Modélisation	24
2.1.2.1 Etapes de la modélisation	24

2.1.2.2 Principe de l'algorithme GOFr	25
2.1.2.3 Construction de la bibliothèque de gaussiennes B	27
2.2 Résultats de la modélisation	30
2.2.1. Prétraitement	31
2.2.2 Résultats graphiques de la bibliothèque de gaussiennes	34
2.2.3. Sélection de la Gaussienne	36
2.2.4. Optimisation paramètre par paramètre	36
2.2.5 Orthogonalisation	38
2.3 Résultats justifiant le choix du nombre de points d'échantillonnage Ne	42
Conclusion	46
Chapitre3 Compression du signal cardiaque	
Introduction	47
<i>Partie A : La transformée en ondelette (wavelet transform : WT)</i>	47
3.1 Passage de la transformée de Fourier à la transformée en ondelette	48
3.2 La théorie des ondelettes	48
3.2.1 Historique	49
3.2.2 Aspect mathématique de la transformée en ondelette	49
3.2.3 La transformée en ondelette continue CWT (Continu wavelet transform)	49
3.2.4 La transformée en ondelette discrète (Discret Wavelet Transform : DWT)	52
3.2.4.1 Expression de la famille d'ondelettes discrètes	52
3.2.4.2 Implémentation de la transformée en ondelette discrète	52
3.2.5. L'orthogonalité	54
3.2.6 Application de la décomposition par DWT à un signal ECG	55
3.2.7 Application de la CWT et de la DWT au signal ECG	57
<i>Partie B : Choix de l'ondelette</i>	61
3.3 Condition de reconstitution du signal	61
3.4 Décomposition et reconstitution pratique du signal ECG	62
<i>Partie C : Compression et reconstitution du signal ECG par ondelette</i>	67
3.5 La notion de compression	67
3.5.1 Compression sans perte	68
3.5.2 Compression avec perte	68
3.6 Principe de la compression	68
3.6.1 Le seuillage	69

3.6.1.1 Types de seuillage	69
3.6.2 Application de la compression au signal ECG et reconstitution	71
3.7 Résultats pratiques et discussion	71
3.7.1 Etude du niveau de décomposition	72
3.7.2 Etude de la valeur du seuil	75
3.8 Tests et résultats	78
Conclusion	82
Chapitre4 Classification, tests et résultats	
Introduction	83
4.1 Analyse en composante principale (ACP)	84
4.1.1. Transformations de l'échantillon –Matrice normalisée	84
4.1.2. Calcul de covariance et de la corrélation	86
4.1.3 Projection	87
4.1.4. Diagonalisation	87
4.1.5 Choix de la dimension de projection	88
4.2 Analyse en Composantes Indépendantes (ACI)	88
4.3 Classification	89
4.3.1 Classification par machines à vecteurs support	90
4.4 Tests et résultats de la classification	93
4.4.1 Résultats de classification moyennant RBF, GOFR et SVM	94
4.4.1.1 Extraction de paramètres temporels moyennant RBF et GOFR	95
4.4.1.2 Classification des signaux ECG par SVM	95
4.4.2 Résultats de classification moyennant DWT, ACP, ACI et SVM	101
4.4.2.1 Extraction de paramètres statistiques moyennant la DWT	101
4.4.2.2 Réduction de dimension	109
4.4.2.3 Apprentissage et Classification	111
4.4.3 Temps de calcul des différentes étapes de décision	118
Conclusion	121
Conclusion générale	123
Bibliographie	126
Annexes	A-3

Liste des tableaux

Tableau1.1: Description de quelques arythmies cardiaques.....	20
Tableau 3.1: Tableau récapitulatif des noms des ondelettes utilisées.....	58
Tableau3.2: Ondelettes les plus performantes pour l'ECG 103.....	65
Tableau3.3: Ondelettes les plus performantes de l'ECG 105.....	66
Tableau3.4: Ondelettes les plus performantes de l'ECG 217.....	66
Tableau3.5: Résultats pour l'ECG 16273.....	79
Tableau3.6 : Résultats pour l'ECG cu01.....	79
Tableau3.7 : Résultats pour l'ECG X_04936.....	80
Tableau3.8 : Résultats pour l'ECG 105	81
Tableau 4.1: Matrice de confusion.....	92
Tableau4.2: Classes de distribution des échantillons pour l'apprentissage et le test.....	98
Tableau4.3: Matrice de de confusion de l'ensemble d'apprentissage.....	98
Tableau4.4: Matrice de de confusion de l'ensemble de Test.....	99
Tableau4.5: Performance du classifieur proposé	99
Tableau4.6 : Performances de l'approche temporelle	100
Tableau4.7 : Bandes de fréquences des différents niveaux de décomposition	103
Tableau4.8 : Distribution des échantillons pour l'apprentissage et le test	112
Tableau 4.9: Performances de classification par SVM de quelques ondelettes.....	114
Tableau4.10: Performances de l'approche temps-Fréquence.....	117
Tableau4.11: Comparaison entre les deux approches.....	119
Tableau4.12: Résultats de comparaison de quelques méthodes de classification des battements cardiaques.....	120

Abréviations

Acc	Accuracy (Taux de classification)
ACP	Analyse en composantes principales
ACI	Analyse en composantes indépendantes
ANN	Artificial Neural Network (réseaux de neurones artificiels)
AV	Auriculo-ventriculaire
aVR	a (augmenté), V (voltage), R (right arm).
aVL	a (augmenté), V (voltage), L (left arm).
aVF	a (augmenté), V (voltage), F (foot).
BD	distribution de Bessel
BIH	Hôpital Beth Israël de Boston
CWT	Continu Wavelet Transform ; transformée en ondelette continue
DCT	transformée en cosinus discrète
DWT	Discret Wavelet Transform ; transformée en ondelette discrète
ECG	Electrocardiogramme
EQM	Erreur Quadratique Moyenne
ESV	Extrasystole ventriculaire
FFT	Fast Fourier Transform ; transformée de Fourier rapide
FHyb-HOSA	les réseaux de neurones hybrides flous
FIR	Finite Impulse Response (Réponse Impulsionnelle Finie)
FN	False Negative (faux négatifs)
FP	False Positive (faux positifs)
GOFR	Generalized Orthogonal Forward Regression
K-NN	K-Nearest Neighbors
LDA	Linear Discriminate Analysis
MLP	Multi Layer Perceptron
MIT	Institut de technologie du Massachusetts

NNM	Neural Network Model
OFR	Orthogonal Forward Regression
PC	Performance de Compression
PNN	probabilistic Neural Network
PR	Performance ou norme de Reconstitution
RBF	Radial Basis Functions
RBFNN	Radial Basis Function Neural Network)
RNN	Recurrent Neural Network
SOM	Self Organizing Map
SPWVD	pseudo distribution de Wigner-Ville lissée
STF	Short Fourier Transform
STFT	Short Term Fourier Transform
SVD	Décomposition en valeurs singulières
SVM	Support Vector Machine
Tc	Taux de classification
TF	Transformée de Fourier
TN	True Negative (Vrais Négatifs)
TNR	True Négatif Ratio (taux de vrais négatifs)
TP	True Positive (Vrais Positifs)
TPR	True Positif Ratio (taux de vrais positifs)
TV	Tachycardie ventriculaire
TWV	distribution de Wigner-Ville
WT	Wavelet Transform (transformée en ondelette

INTRODUCTION GENERALE

Les maladies cardiovasculaires constituent un problème majeur de santé publique et l'une des causes de mortalité des plus fréquentes dans notre pays. En raison de l'ampleur du problème, le diagnostic précoce des anomalies cardiovasculaires, la prévention et le suivi des patients à risques, devient une nécessité. L'électrocardiogramme (ECG) est l'examen, le moins coûteux, non invasif, le plus rapide, qui permet la détection des anomalies cardiaques. Il est très souvent complété, par un examen similaire d'une durée de 24 heures appelé "Holter", examen au cours duquel le patient peut vaquer à ses occupations habituelles. L'avantage du Holter est qu'il permet la détection d'évènements sporadiques qui n'interviennent pas nécessairement au cours des quelques secondes de l'enregistrement ECG lorsque le patient est au repos. Avec une surveillance permanente c'est-à-dire le monitoring ambulatoire de l'électrocardiogramme (ECG) du signal cardiaque, il est possible de prévenir une crise cardiaque sans nécessiter une prise en charge hospitalière lourde et coûteuse.

Nos travaux de thèse sur la modélisation et compression du signal cardiaque en vue d'une détection automatique des anomalies cardiovasculaires englobent trois phases:

- ❖ Lecture automatique du signal "Holter" et détection des battements cardiaques de l'ECG
- ❖ Elaboration et implémentation de méthodes de modélisation et de compression en vue d'une réduction de l'espace de mesure, pour un apprentissage automatique.
- ❖ Réalisation d'un système expert pour la détection automatique des anomalies cardiovasculaires.

Phase I :

Le signal ECG doit subir des prétraitements en vue d'une modélisation. Notre travail a porté d'abord sur l'élimination de la ligne de base, en utilisant des filtres numériques. Les résultats obtenus nous ont permis de ramener l'ECG sur une ligne horizontale correspondant au zéro pour permettre une comparaison simple des amplitudes des ondes d'un battement à un autre. Pour transmettre le signal ECG ou réaliser un dépistage automatique, le traitement se fera par battement. Il est donc impératif de faire un découpage de l'enregistrement des signaux cardiaques. Dans le processus de découpage, la détection des complexes QRS est une étape incontournable. Elle nous a permis le calcul des segments RR et de définir ainsi chaque battement du signal. Néanmoins, en vue d'éventuels travaux futurs telle que la modélisation, nous avons préféré

procéder à un découpage à partir de l'onde T afin d'avoir toutes les ondes présentes dans un même battement. Chaque battement s'étend de la fin de l'onde T à la fin de l'onde T du battement suivant.

Phase II :

Les objectifs essentiels de la modélisation du signal ECG sont :

- la mise en évidence des caractéristiques temporelles, énergétiques et fréquentielles du signal ECG.
- la description de son spectre par un ensemble très limité de paramètres.
- l'utilisation des paramètres en télémédecine pour l'enregistrement, le stockage et la transmission des signaux ECG.
- l'utilisation des paramètres pour contribuer directement à une automatisation de diagnostic en résolvant les problèmes de classification.

Il s'agit donc de trouver un modèle mathématique optimal pour représenter un ECG, tout en essayant d'atteindre une erreur minimale entre le signal réel et le modèle.

Pour résoudre ce problème, nous nous sommes intéressés à deux méthodes mathématiques puissantes pour la modélisation du signal cardiaque, qui sont les réseaux de neurones et plus particulièrement les réseaux à bases radiales (RBF) et les ondelettes.

- La modélisation par RBF nous intéresse d'autant plus qu'elle se base sur une fonction gaussienne. Le modèle obtenu est la décomposition du signal ECG en une somme de fonctions gaussiennes. Ce qui permettra d'associer à chaque onde caractéristique de l'ECG une fonction gaussienne. Les paramètres représentant les moyennes et les écarts types et les amplitudes des gaussiennes, seront initialisés et optimisés pour minimiser l'erreur quadratique moyenne appelée coût J. Optimiser les paramètres revient à adapter les paramètres des gaussiennes sélectionnées au tronçon de l'électrocardiogramme à modéliser. Le principe et le formalisme de la modélisation par RBF seront détaillés en chapitre deux.

- la compression demeure un outil très utile pour la transmission en vue d'une analyse ultérieure, d'un diagnostic assisté par ordinateur, d'une constitution d'une banque de données électrocardiographiques, et dans les systèmes de traitement ou de transmission en temps réel. La méthode de compression que nous avons utilisée est la transformée en ondelettes discrète DWT (Discret Wavelet Transform) qui fait usage des caractéristiques du domaine temporel et fréquentiel d'un signal. Elle permet d'obtenir les sous signaux par application récursive, d'abord sur le signal source puis sur le signal d'approximation obtenu à chaque niveau, d'un banc de filtres passe-bas-passe haut. Les performances des méthodes de compression par ondelettes dépendent de

l'ondelette mère, du nombre de niveaux de décomposition, des approches de seuillage et de quantification des coefficients d'ondelettes. Elles sont évaluées en fonction de deux critères :

- Le taux de compression TC qui chiffre de degré de compression obtenu.
- La distorsion qui est l'erreur introduite par l'opération de compression entre le signal reconstruit et le signal original. Elle est exprimée par la mesure de la racine carrée de l'erreur en pourcentage PRD (the percent root mean square difference).

Nous avons effectué plusieurs essais pour choisir l'ondelette qui donne le meilleur compromis entre le PRD et le TC. Les performances de ces algorithmes sont évidemment évaluées par les courbes CR-PRD représentées dans le chapitre III. La compression par l'ondelette qui donne un meilleur taux de compression avec un minimum de PRD garantit la qualité acceptable d'une manière clinique et aussi la conservation de l'information considérable après reconstruction (CR, PRD et observation visuelle du clinicien). Nous avons pu démontrer que l'ondelette bi orthogonale est la mieux adaptée à la modélisation des signaux ECG sains et pathologiques. De bons taux de compression et une assez bonne reconstitution sont obtenus. En terme de reconstitution du signal ECG, cette méthode est intéressante car elle contribue à sa transmission. Pour le diagnostic, l'extraction des paramètres caractéristiques des signaux n'est autre que celle des coefficients de détails d'ondelettes.

Phase III :

La détection automatique des anomalies cardiovasculaires repose sur :

- ✓ les résultats des deux modélisations précédentes à savoir : l'extraction d'informations caractéristiques dans le domaine temporel (le signal ECG considéré comme une fonction du temps) et dans le domaine fréquentiel, par décomposition multi résolution.
- ✓ Les méthodes de classification utilisées : le classificateur proposé utilise une base de données comportant deux types de signaux: sains et pathologiques. Il a pour but d'identifier 2 types de battements cardiaques dénommés: Normal et Anormal.

Nous nous sommes fixés comme objectif principal la conception d'un système d'aide au diagnostic permettant, à partir d'enregistrements à long terme du signal ECG (Holter) contenant des centaines de milliers de battements de cœur, de faire l'analyse de ces derniers, leur identification, l'apprentissage automatique et donc la détection de deux classes : signaux normaux et pathologiques, moyennant des outils mathématiques robustes.

Notre travail est fondé sur un ensemble de méthodes qui permettent la conception d'un système en temps réel qui réalise à partir d'un signal biomédical ECG une extraction en ligne des informations qui vont permettre d'établir une aide au diagnostic médical.

Le système aura pour rôle de choisir les éléments à extraire du signal biomédical, décrivant ainsi ce signal dans un langage plus ou moins précis, et sélectionner le langage optimal à utiliser pour établir le diagnostic en ligne. Nous avons entamé cet axe de recherche en étudiant et en développant des algorithmes appropriés utiles pour chaque phase de traitement de l'ECG.

La conception de notre système d'aide au diagnostic des signaux biomédicaux ECG repose dans une première étape sur l'extraction de paramètres pertinents à partir d'ECG réels, l'apprentissage puis leur classification en signaux normaux et anormaux dans la seconde étape.

Pour ce faire, nous proposons deux approches d'extraction de paramètres. La première approche utilise une méthode temporelle basée sur la modélisation du signal ECG. La seconde qui est fréquentielle, basée sur la décomposition temps-fréquence du même signal, s'est avérée intéressante pour la compression de l'espace de mesure.

La méthode temporelle utilise les réseaux de neurones RBF (réseaux à bases radiales) et l'algorithme GOFR (Generalized Orthogonal Forward Regression) pour modéliser chaque battement du signal cardiaque, par cinq gaussiennes avec des amplitudes, écarts types et moyennes différents.

La méthode temps-fréquence qui repose sur la transformée en ondelettes discrètes DWT dont la décomposition en bandes de fréquences recouvre les caractéristiques fréquentielles des différentes ondes de l'ECG (P, QRS, T), a permis d'exploiter les cinq premiers niveaux de décomposition. La combinaison des paramètres temporels issus de la modélisation par RBF et GOFR avec l'algorithme de classification SVM et les paramètres fréquentiels issus de la décomposition par DWT et de la réduction de l'espace de mesure par l'ACP ou l'ACI avec le même algorithme de classification SVM, a permis d'atteindre des taux de reconnaissance de classification très importants.

Dans notre travail, nous montrons qu'il est possible de parvenir à une meilleure performance de classification pour les signaux ECG, avec une seule variété de paramètres statistiques: la variance de l'énergie.

Les signaux ECG ont été décomposés en représentations temps-fréquence à l'aide de la DWT et la variance de l'énergie (fonction statistique) des coefficients de détail des cinq premiers niveaux de décomposition, a été calculée pour représenter leur distribution. La méthode de réduction de paramètres ACP-ICA a été combinée avec les SVM pour parvenir à une classification des signaux en normaux ou anormaux, en réduisant le temps de calcul. Les comparaisons sont fondées sur deux mesures de la performance issues des matrices de confusion: la spécificité et la sensibilité. Le résultat de la classification du signal ECG en utilisant SVM montre que l'extraction de caractéristiques non linéaire peut améliorer la performance du système de classification.

Dans notre thèse, le premier chapitre sera consacré au système cardiovasculaire où l'anatomie du cœur, les ondes de l'ECG, leurs significations pratiques et certaines arythmies seront résumées. Le second chapitre développera une technique de modélisation temporelle et d'extraction de paramètres pertinents par les réseaux de neurones RBF. Le troisième chapitre abordera la compression avec reconstitution des signaux ECG moyennant une méthode temps-fréquence: la théorie des ondelettes discrètes DWT. Dans le quatrième chapitre la classification par utilisation des support vecteur machine SVM sera appliquée aux signaux ECG pour les deux techniques de modélisations temporelle et fréquentielle, en se basant sur les méthodes d'analyses par composantes principales ACP et indépendantes ACI. Les résultats de classification en vue d'un apprentissage automatique et de détection automatique des ECG normaux et anormaux seront discutés et interprétés.

En conclusion, nous soulignerons l'essentiel de notre contribution et proposerons quelques perspectives.

ETAT DE L'ART

De nombreuses recherches et techniques ont été développées pour analyser le signal ECG en vue d'un diagnostic automatique. Les systèmes proposés ont été principalement basés sur les étapes suivantes :

- prétraitement
- extraction de paramètres
- réduction de dimension
- classification

La précision de l'acquisition du signal est d'une grande importance car elle contribue de façon significative au résultat global de classification, d'où la nécessité d'un prétraitement.

Le prétraitement consiste à filtrer les signaux ECG pour les débarrasser de leurs artefacts en utilisant une succession de filtres numériques [1].

Les caractéristiques extraites du signal ECG jouent un rôle essentiel dans le diagnostic de la maladie cardiaque.

Elles peuvent être extraites dans le domaine temporel [2] ou dans le domaine fréquentiel [3].

Dans le domaine temporel le système d'extraction de paramètres de l'ECG fournit les caractéristiques fondamentales (amplitudes et intervalles de temps) pour une analyse automatique ultérieure. Dans [4,5], l'ensemble de caractéristiques obtenu est basé sur la morphologie de l'ECG et les intervalles RR et ST.

une méthode pour l'extraction automatique à savoir les intervalles de temps et les caractéristiques morphologiques, de l'électrocardiogramme (ECG) pour classer les ECG en normal et arythmique a été décrit dans [6,7 ,8]. La méthode utilisée est la combinaison de réseaux de neurones artificiels (ANN) et l'analyse discriminante linéaire (LDA). Un total de 10 caractéristiques à savoir RR, RTC, amplitude de l'onde T, l'inclinaison de l'onde T, l'aplatissement de l'onde T, la durée des ondes a été utilisé.

L'analyse par les fonctions de corrélation [9] pour l'extraction de caractéristiques du signal ECG anormal, est une technique d'analyse basée sur l'analyse de corrélation croisée. Elle mesure la similarité entre deux signaux et extrait les informations présentes dans ces signaux. Les résultats de cette technique montrent qu'il est possible d'extraire des caractéristiques qui différencient les deux types de signaux cardiaques : normaux et anormaux.

Un algorithme [10] pour la détection des complexes QRS, utilise une définition modifiée de la pente du signal ECG, comme caractéristique de détection du QRS. Une succession de transformations : filtrage, correction de la ligne de base, sont réalisées au préalable.

L'extraction de caractéristiques statistiques à partir des signaux bruts ou bien de leur décomposition en ses sous-bandes caractéristiques utilise les techniques de transformation temps-fréquence comme les transformées en ondelettes continues (CWT), transformées en ondelettes discrètes (DWT) [11], transformée en cosinus discrète (DCT) [12], la distribution de Bessel (BD), la distribution de Wigner-Ville (TWV), pseudo distribution de Wigner-Ville lissée (SPWVD) [13], Lyapunov exposants [14], l'entropie [15] de Renyi

Les caractéristiques peuvent être de domaine spectral ou temporel comme les énergies, les moyennes, les écarts types ou médiane, maximum, minimum, la variance écart absolu moyen... etc [16].

La réduction de la dimension de l'espace de mesure, permet d'éviter la redondance, de garder uniquement les paramètres pertinents et donc de diminuer le coût de calcul, et ce par l'utilisation de l'analyse en composantes principales (ACP) [17.], de l'analyse en composantes indépendantes (ICA) [18,19,20] ou de l'analyse discriminante linéaire (LDA) [6].

La dernière étape est la classification binaire des signaux ECG en signaux normaux ou anormaux en utilisant un classificateur approprié. Il ya un grand choix de classificateurs comme les support vector machines (SVM), SVM floue, les différents réseaux de neurones

artificiels (probabiliste NN, feed-forward NN, perceptron multi-couche NN (MLP)...) [6], l'algorithme de classification de vecteur de quantification linéaire (LQV) [21] ... etc

Toutes ces techniques et algorithmes proposés dans la littérature pour extraire les caractéristiques d'un signal ECG ont leurs avantages et leurs limites.

Les critères qui doivent être considérés lors de l'élaboration d'un algorithme pour l'extraction de caractéristiques d'un signal ECG sont la simplicité de l'algorithme et sa précision en fournissant les meilleurs résultats dans l'extraction de caractéristiques. Dans les divers travaux énumérés, les résultats expérimentaux ont montré que le taux de reconnaissance a pu atteindre 99% et plus. Cependant le nombre de paramètres caractéristiques mis en jeu est au minimum égal à 15 et plus.

Dans le souci de réduire le nombre de paramètres et donc le temps de calcul et d'atteindre un taux de reconnaissance élevé, nous avons proposé deux combinaisons de méthodes à savoir : RBF-GOFR-SVM et DWT-ACI-SVM.

La méthode temporelle utilise les réseaux de neurones RBF (réseaux à bases radiales) et l'algorithme GOFR (Generalized Orthogonal Forward Regression) pour modéliser chaque battement du signal cardiaque par cinq gaussiennes avec des amplitudes, écarts types et moyennes différents.

La méthode temps-fréquence qui repose sur la transformée en ondelettes discrète DWT dont la décomposition en bandes de fréquences recouvre les caractéristiques fréquentielles des différentes ondes de l'ECG (P, QRS, T), a permis l'extraction de cinq premiers coefficients de détails. L'application de l'ACP/ACI à la variance énergétique de ces coefficients a réduit l'espace de mesure.

Pour un même classifieur SVM, l'approche temporelle basée sur la modélisation RBF combinée à l'algorithme GOFR, a permis d'atteindre un taux maximal de reconnaissance de 99.67% alors que l'approche fréquentielle associée à l'algorithme de l'ACI a permis d'atteindre un taux de 100%.

Introduction

Dans ce chapitre, nous allons présenter le système cardiovasculaire, le principe de l'électrocardiogramme l'ECG d'une manière succincte, puis énumérer les principales ondes et paramètres caractérisant un signal ECG sain d'un ECG pathologique.

1.1 Le cœur

Le cœur (figure 1.1) est l'élément central du système cardiovasculaire, [22-25].

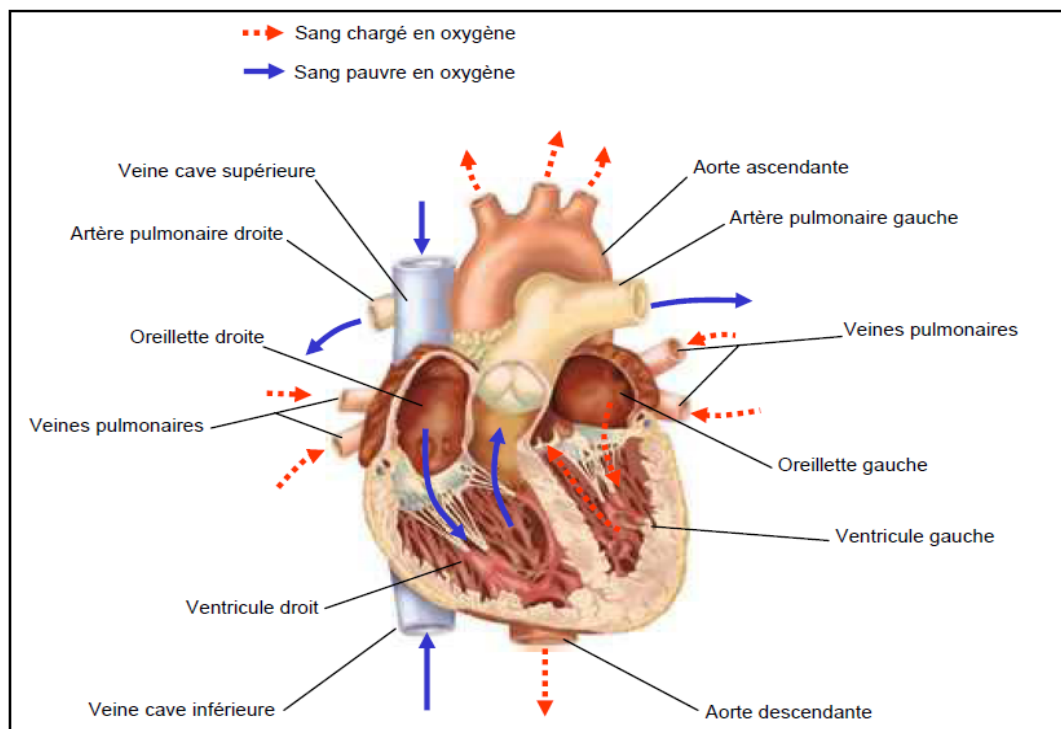


Figure 1.1: Coupe du cœur [24].

Le cœur est séparé en deux : le cœur gauche et le cœur droit, par une épaisse cloison. Chaque côté est composé de deux cavités : une oreillette et un ventricule reliés entre eux par une valve à sens unique. L'oreillette droite récupère le sang pauvre en oxygène (flèches bleues continues) de la veine cave supérieure et inférieure, et le propulse par contraction de son tissu dans le ventricule droit. A son tour il se contracte et envoie le sang dans les poumons où il est chargé en oxygène. Le sang revient au cœur dans l'oreillette gauche, puis passe dans le ventricule gauche et est envoyé vers les organes dans le réseau artériel (flèches rouges en pointillés).

1.2 Fonctionnement électrique

Comme pour tous les muscles du corps, la contraction du myocarde est provoquée par la propagation d'une impulsion électrique le long des fibres musculaires cardiaques induite par la dépolarisation des cellules musculaires.

Le cœur effectue un cycle cardiaque (figure 1.2) comportant trois phases qui sont: **la systole auriculaire** où les oreillettes se contractent et poussent le sang vers les ventricules, **la systole ventriculaire** où la contraction des ventricules pousse le sang vers les artères **et la diastole**, qui est la phase de relâchement générale, elle permet le remplissage des ventricules par les oreillettes depuis les veines cave et pulmonaire.

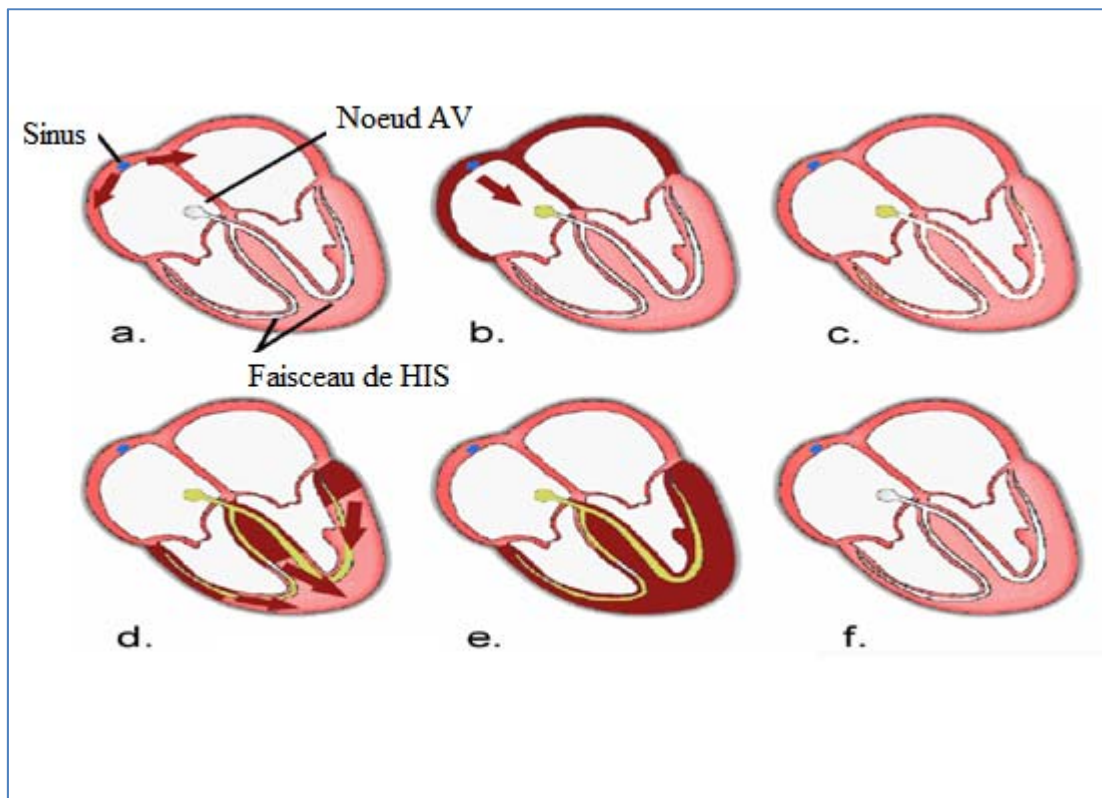


Figure 1.2: Le cycle cardiaque[22].

L'impulsion électrique induite par la dépolarisation des cellules, se propage dans le muscle cardiaque et induit sa contraction. Elle prend naissance dans le sinus (a) puis se propage dans les oreillettes (b) entraînant leurs contractions (systole auriculaire). L'impulsion arrive alors au nœud auriculo-ventriculaire (AV), seul point de passage électrique entre les oreillettes et les

ventricules. Une courte pause est alors introduite (c) juste avant la propagation dans les fibres constituant le faisceau de His. Au passage de l'impulsion électrique (d), les ventricules se contractent à leur tour (e) (systole ventriculaire). Après la diastole (décontraction du muscle) les cellules se repolarisent (f). Le cycle du battement cardiaque est alors terminé et le cœur est prêt pour un nouveau battement, [22-25].

1.3 Principe de l'électrocardiogramme (ECG)

En cardiologie, l'examen le plus couramment pratiqué est l'électrocardiogramme (ECG), [26,27].

En 1901, Willem Einthoven utilise un galvanomètre à cordes, le courant mesuré par des électrodes sur le torse du patient mettait en mouvement un mince fil d'argent tendu entre les pôles d'un gros aimant; ces déflexions étaient enregistrées sur du papier photographique qui se déroulait face à un rayon lumineux. Ce qui a permis d'enregistrer l'activité électrique du cœur anormale et la comparer à l'activité normale. Ainsi naît l'électrocardiogramme. Il publie les premières classifications d'électrocardiogrammes pathologiques en 1906. Ses travaux sur l'électrocardiographie lui vaudront le prix Nobel physiologie ou médecine en 1924.

En 1942, Emanuel Goldberger réalise le premier électrocardiogramme sur 12 voies, qui est toujours utilisé. Le principe de l'enregistrement moderne est, à peu de chose près, celui qui fut proposé par Einthoven : grâce à deux électrodes collées à la surface de la peau, on enregistre la différence de potentiel entre deux points diamétralement opposés par rapport au cœur, ce signal étant directement corrélé au déplacement de l'impulsion électrique dans les fibres du muscle cardiaque.

L'activité électrique instantanée peut être définie par un vecteur orienté suivant la différence de potentiel présente dans le cœur, et de module proportionnel à celle-ci. Le couple d'électrodes enregistre à chaque instant l'amplitude de la projection de ce vecteur suivant leur axe: ainsi, lorsque le vecteur électrique est orienté de l'électrode - à l'électrode +, on observe sur l'enregistreur une déflexion positive, et lorsque le vecteur est orienté en sens inverse, la déflexion est négative.

C'est l'enregistrement de l'activité électrique du cœur en fonction du temps, moyennant des électrodes cutanées placées selon des dérivations bien définies, avec une vitesse de déroulement du papier à 25 mm par seconde et une amplitude de 10 mm pour 1 mV. Ce standard a été adopté

il y a plus de 50 ans, figure 1.3. Ces dérivations permettent d'avoir une idée sur l'activité électrique du cœur.

Le signal recueilli n'est autre qu'une projection de l'axe électrique du cœur selon la direction de la dérivation (comme perçu par les électrodes). L'emplacement des électrodes est en général choisi de sorte à explorer la quasi-totalité du champ électrique cardiaque en offrant un ensemble cohérent de dérivations non redondantes. Plusieurs systèmes standardisés existent.

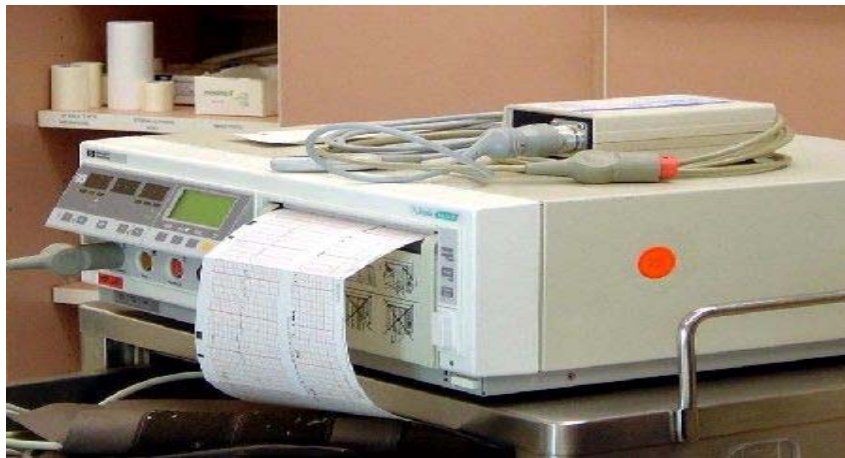


Figure 1.3: L'électrocardiographe.

L'ECG est enregistré selon 12 dérivations. Il est visualisé selon 12 axes privilégiés : 6 dans le plan frontal (les dérivations des membres) : D1, D2, D3, aVR, aVL, aVF, et 6 dans le plan horizontal (les dérivations précordiales) : V1 à V6, figure I.4.

1.3.1 Dérivations des membres

Les dérivations des membres sont placées à la face interne des poignets et des chevilles. Elles explorent le champ électrique cardiaque dans un plan frontal. Les électrodes du plan frontal forment des dérivations : bipolaires ou unipolaires.

• bipolaires des membres ou « standard » :

DI : bras droit (-), bras gauche (+)

DII : bras droit (-), jambe gauche (+)

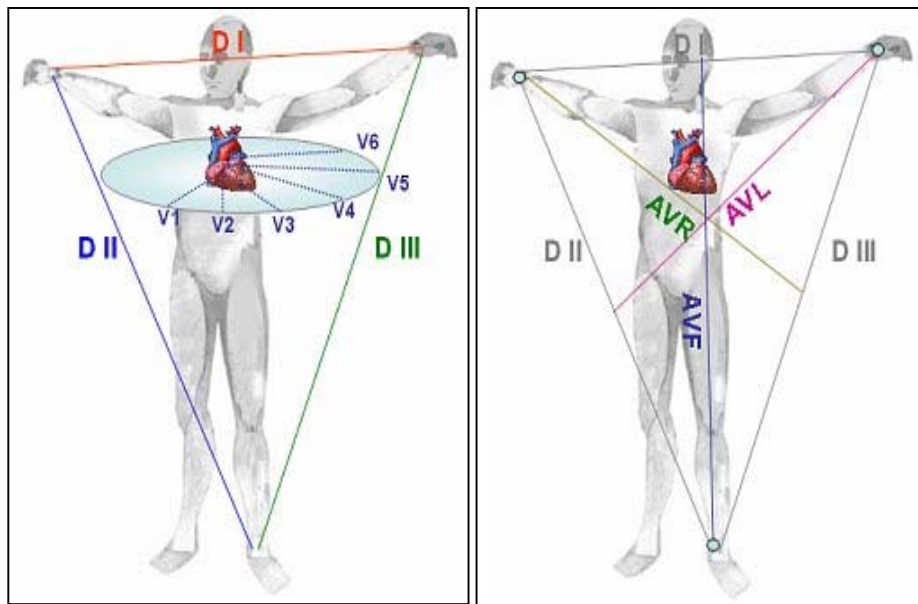


Figure 1.4: Les douze dérivations de l'ECG.

DIII : bras gauche (-), jambe gauche (+).

Ces trois dérivations forment le triangle équilatéral d'Einthoven, le montage des polarités étant tel que $DII = DI + DIII$ (figure I.5)

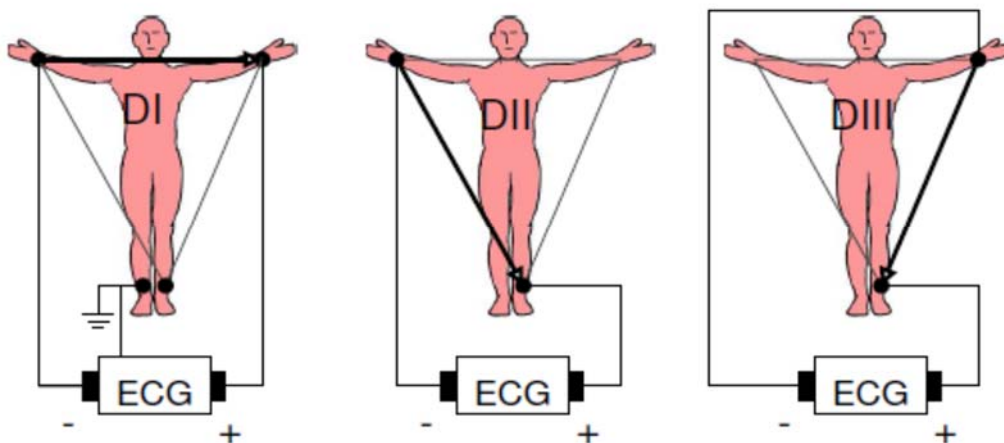


Figure 1.5: Dérivations bipolaires des membres.

•**unipolaires :**

L'électrode exploratrice positive est l'un des membres, les autres étant reliées à une borne centrale de potentiel nul. Pour obtenir un tracé d'amplitude similaire aux trois dérivations bipolaires, il faut les amplifier, (figure 1.6) d'où leurs dénominations :

aVR : a (augmenté), V (voltage), R (right arm).

aVL : a (augmenté), V (voltage), L (left arm).

aVF : a (augmenté), V (voltage), F (foot).

Ces six dérivations étant dans le même plan frontal, la translation de leurs axes au centre du triangle d'Einthoven permet de construire un système de coordonnées (double triaxe de Bailey), utile au calcul de l'axe du vecteur d'activation dans le plan frontal .

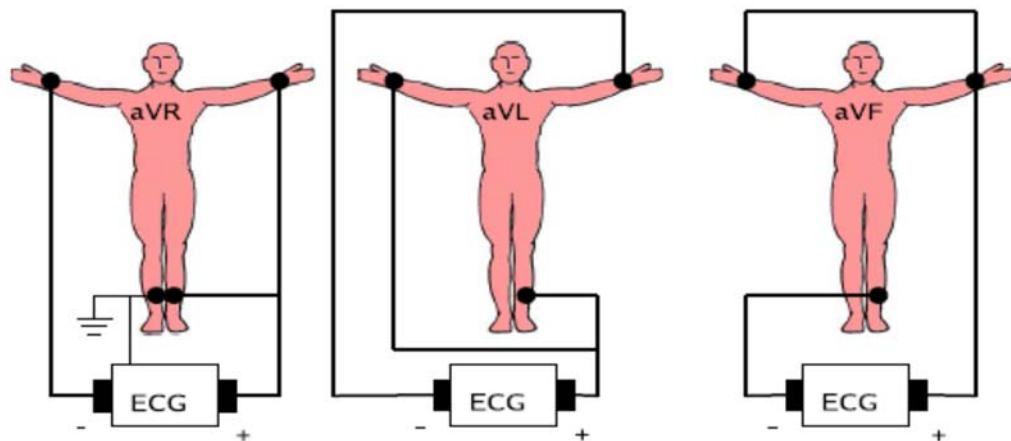


Figure 1.6: Dérivations unipolaires des membres.

1.3.2 Dérivations précordiales :

Les dérivations précordiales sont unipolaires et explorent l'activité électrique cardiaque dans le plan horizontal. La position de chaque électrode sur le thorax doit être précise (figure 1.7) pour permettre la comparaison entre des ECG successifs :

- V1 : 4ème espace intercostal droit, au ras du sternum (on ne compte pas l'espace entre la clavicule et la première côte comme un espace intercostal).
- V2 : symétrique par rapport au sternum (4ème espace intercostal gauche).
- V3 : à mi-distance entre V2 et V4.

- V4 : 5ème espace intercostal gauche, sur la ligne médio-claviculaire.
- V5 : sur la ligne axillaire antérieure à « l'horizontale » de V4.
- V6 : sur la ligne axillaire moyenne à « l'horizontale » de V4.

Il peut être utile d'ajouter :

- V3 R, V4 R : symétriques, à droite de V3-V4.

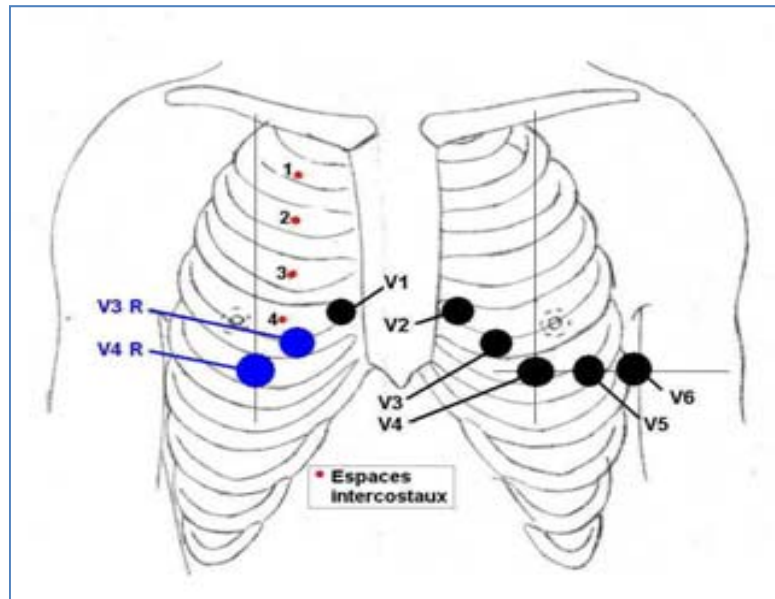


Figure 1.7: Dérivations précordiales.

L'ECG permet le diagnostic et la localisation précise de certaines pathologies qui laissent des traces permanentes. La durée d'un ECG standard peut varier de quelques secondes à une ou deux minutes, [26,27].

1.4 L'ECG de Holter

Au début des années 60 Norman Holter a remarqué que la courte durée de l'examen ECG 12 est un obstacle à la détection systématique de pathologies qui apparaissent de manière sporadique, comme certains troubles du rythme par exemple.

C'est pourquoi il proposa un appareil « portatif I » permettant d'enregistrer l'activité cardiaque pendant plusieurs heures ; cet enregistrement constitue ce que l'on appelle « l'examen Holter », [28,29].

L'enregistreur Holter d'aujourd'hui (Figure 1.8) , se présente sous la forme d'un petit boîtier auquel sont reliées sept électrodes, permettant ainsi l'enregistrement sur 2 ou 3 dérivations (2frontales et 1 transversale). Le patient se fait poser l'appareil chez un cardiologue et retourne ensuite à ses occupations habituelles. 24 heures plus tard, il revient chez le cardiologue pour se faire enlever l'appareil qui a en mémoire 24 heures d'enregistrements ECG effectués tandis que le patient a vécu une journée normale. Les résultats issus de l'analyse des quelque 100 000 battements que compte l'enregistrement permettent ainsi de diagnostiquer une plus grande gamme de pathologies que l'ECG hospitalier.



Figure 1.8 : Enregistreur Holter.

L'examen Holter, examen non invasif complémentaire de l'ECG 12 dérivations, est de plus en plus prescrit ; depuis quelque temps. Des enregistrements Holter peuvent être de très longues durée : 1 mois par exemple. Ce qui correspond à plus de 2 millions de battements enregistrés. L'analyse de ces enregistrements n'est envisageable que si une lecture automatique des données enregistrées est possible, [28,29].

1.5 Analyse de l'ECG

L'Analyse des différents tracés électriques du cœur a permis de concevoir un signal ECG standard (Figure 1. 9), [27].

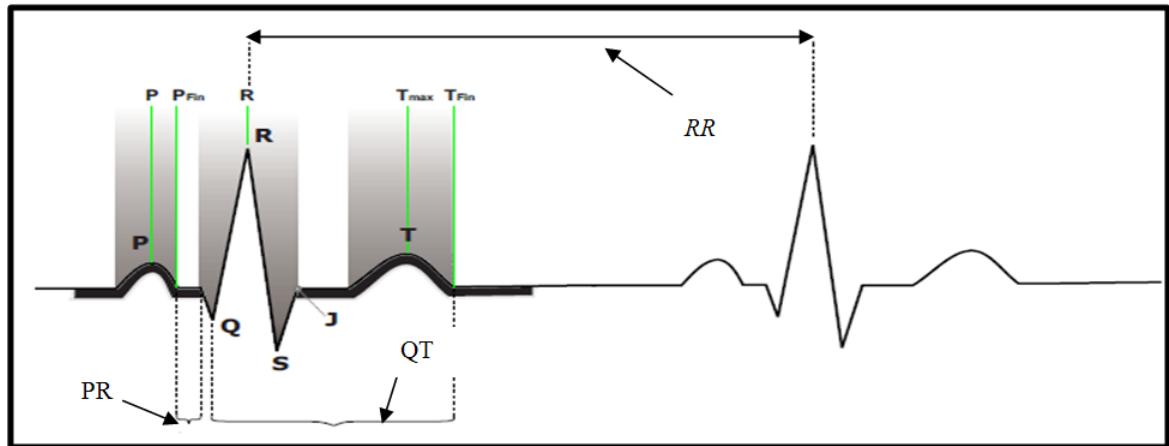


Figure 1.9: Ondes et intervalles de l'ECG [27].

1.5.1 Ondes de l'ECG

Le signal ECG est constitué des principales ondes électriques suivantes :

1.5.1.1 l'onde P

La dépolarisation des oreillettes est perçue comme l'onde P de l'ECG (figure I.10.a). Elle a une durée de 0.1 à 0.12 secondes et une amplitude de 0.1 à 0.3 mV. Elle est positive pour les dérivations DI, DII, DIII, et V6. Le spectre de fréquence de l'onde P est entre 10 et 15 Hz. Chaque onde P doit précéder un complexe QRS.

1.5.1.2 le complexe QRS

La première dépolarisation ventriculaire correspond au complexe QRS (figure 1.10.b), qui est constitué de trois ondes : l'onde R qui est l'onde positive de 0.5 à 2 mv, l'onde Q qui est la première onde négative précédant l'onde R et enfin l'onde S qui est la première onde négative après R. Le spectre de fréquence du complexe QRS se trouve entre 20 et 50 Hz.

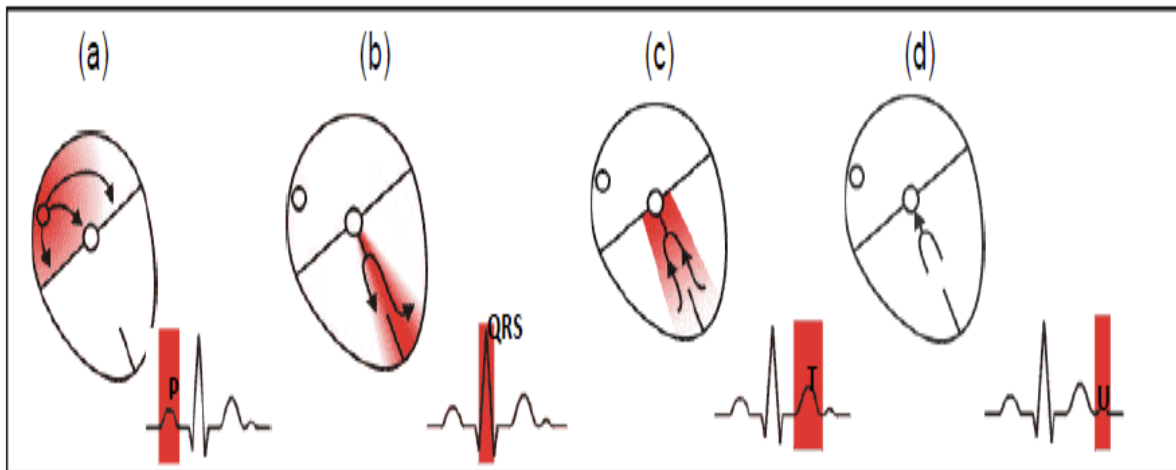


Figure 1.10: Les différentes ondes du tracé ECG.

1.5.1.3 l'onde T

Elle exprime la repolarisation des ventricles (figure 1.10.c). Elle a une durée de 200ms et une amplitude de 0.2 mv. La fréquence de l'onde T est de 5 Hz, elle est toujours positive en DI et DII et V6, négative en aVR.

1.5.1.4 l'onde U

Après l'onde T, un enregistrement ECG peut parfois présenter une onde de très basse amplitude dite l'onde U (figure 1.10.d), elle est similaire dans la forme à l'onde P, et elle est souvent associée aux processus de repolarisation ventriculaire tardive.

1.5.2 Intervalles et segments de l'ECG

On détaille ci-dessous les intervalles et les segments d'un ECG normal (figure 1.11), [30- 32].

1.5.2.1 Intervalle PR

Il mesure la distance entre le début de l'onde P et le début du complexe QRS, il représente la dépolarisation des oreillettes et du nœud AV. Sa durée normale est entre 0.12 et 0.2 secondes.

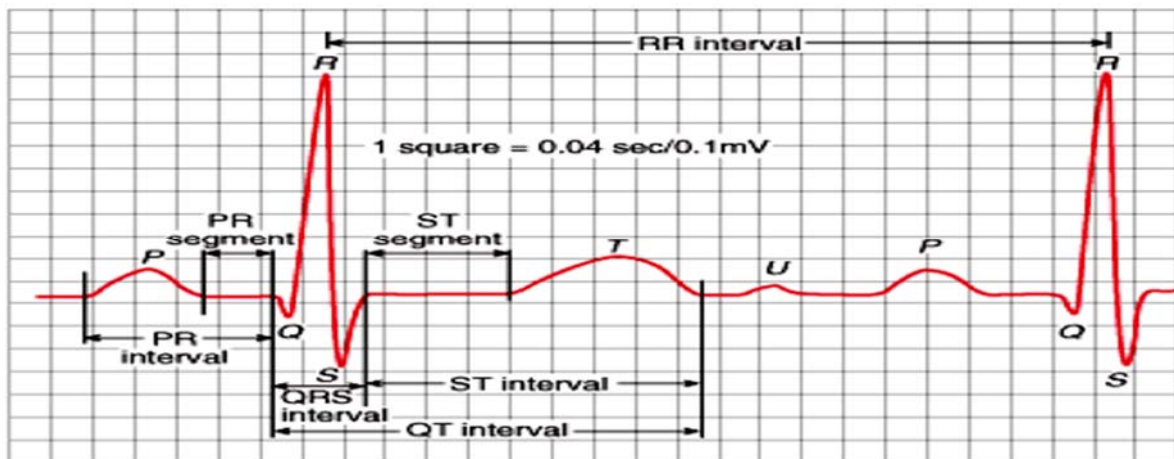


Figure 1.11: Intervalles et segments d'un ECG normal.

1.5.2.2 Segment PR

C'est la période temporelle comprise entre la fin de l'onde P et le début du complexe QRS, il représente le temps de transmission du front de dépolarisation par le nœud auriculo-ventriculaire AV.

1.5.2.3 Segment ST

Il est compris entre la fin du complexe QRS et le début de la phase ascendante de l'onde T. Ce segment correspond au temps pendant lequel l'ensemble des cellules myocardiques sont dépolarisées et donc, dans le cas normal, il doit être isoélectrique. Dans le cas contraire, le niveau d'amplitude et la pente de ce segment sont des indicateurs de l'état ischémique du myocarde.

1.5.2.4 Intervalle QT

C'est le temps entre le début du complexe QRS et la fin de l'onde T. Il représente une indication de la longueur des phases de dépolarisation et repolarisation ventriculaires. Sa durée varie avec la fréquence cardiaque entre 0.3 et 0.38 secondes.

1.5.2.5 Intervalle PP

Il mesure la durée du cycle des oreillettes, c'est un indicateur de la fréquence des oreillettes.

1.5.2.6 Intervalle RR

La distance entre deux ondes R successives, c'est l'indicateur de la fréquence des ventricules.

1.6 Contenu spectral de l'ECG

Le signal ECG est un signal riche par la variété des ondes qui le constitue ondes (P, QRS et T). Ainsi, il présente une densité spectrale de puissance qui varie en fonction de la morphologie du signal (dérivations) et d'un sujet à un autre. Les densités spectrales de puissance obtenues par la méthode de périodogramme moyenné sur 150 battements sont illustrés sur la figure (1.12), [32-33].

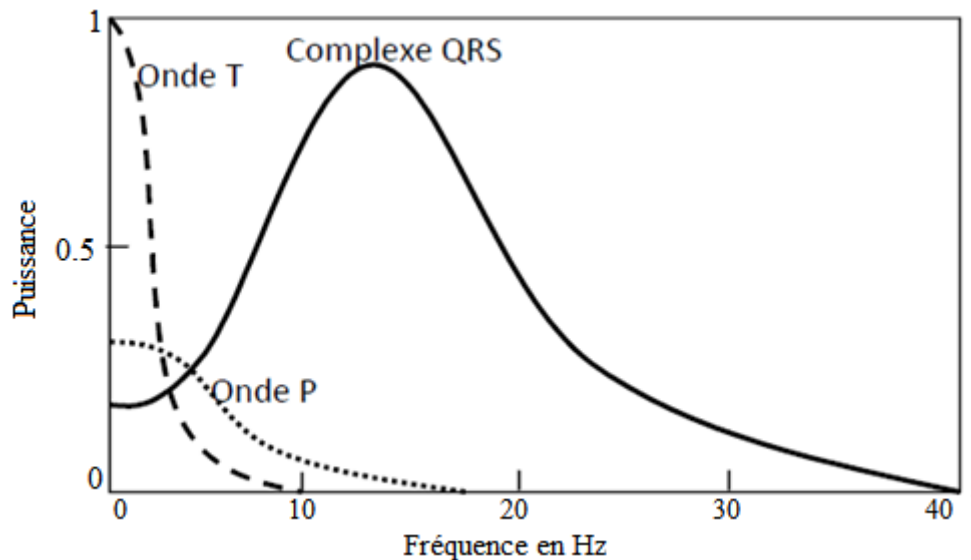


Figure 1.12: Spectre de puissance de l'onde P, du complexe QRS et de l'onde T.

On y observe le comportement passe-bas de l'ECG dont l'énergie est comprise entre 1 et 30Hz. Dans le cas des ondes P et T, le contenu spectral se limite aux très basses fréquences, tandis que pour le complexe QRS, son contenu spectral a une tendance vers les plus hautes fréquences.

I.7 Les arythmies du cœur

L'arythmie survient lorsque l'excitation électrique naît ailleurs que dans le nœud sinusal, dans le nœud atrio-ventriculaire ou les ventricules par exemple, ou que l'onde électrique ne suit plus les voies normales de propagation. La contraction cardiaque qui en résulte s'écarte du rythme normal. On classe alors les arythmies selon leurs effets sur le rythme cardiaque et leur lieu de formation. Si une arythmie accélère le rythme cardiaque, on parle de tachycardies si elle le ralentit, on parle alors de bradycardie. Si le trouble se déclare dans les oreillettes ou le nœud

atrio-ventriculaire, on parle d'arythmie supra ventriculaire; s'il apparait dans les ventricules, c'est l'arythmie ventriculaire, [34- 37].

Il y a plusieurs formes d'arythmies, celles qui touchent les ventricules, sont en général plus graves que les arythmies qui touchent les oreillettes. Nous résumons les arythmies cardiaques dans le tableau1.1 suivant :

Tableau1.1 : Description de quelques arythmies cardiaques

Coordonnées Types	Définition	Caractéristiques	Détails
Extrasystole ventriculaire ESV	Une contraction cardiaque prématurée.	QRS prématuré non précédé d'onde P. Inversion de l'onde T	QRS élargi ≥0,12s
Tachycardie ventriculaire TV	Due à une activation non contrôlée des ventricules.	Fréquence cardiaque >160bat/mn RR est parfois irrégulier	QRS large de 0,12s. -Les ondes P sont dissociées
Torsades de pointes	Dues à un trouble de repolarisation ventriculaire QT allongé : repolarisation retardée.	Fréquence ventriculaire [200-250] Bat/mn	L'amplitude et la polarité de QRS variables.
Fibrillation ventriculaire	Un arrêt cardiovasculaire immédiat avec anoxie cérébrale	L'activité cardiaque est rapide, irrégulière, anarchique.	QRS larges >0,12sec. Onde T anormale.
Extrasystole Auriculaire	Un trouble cardiaque dû à une contraction prématurée d'une des cavités du cœur.	La forme de l'onde P est différente.	QRS peut être normal et parfois anormal.
Flutter auriculaire	Un trouble du rythme cardiaque des battements cardiaques rapides et ininterrompus par des pauses.	Son rythme : 300 Battements/mn	l'oreillette et le ventricule ne sont pas synchronisés
Fibrillation auriculaire	Réentrée continue de l'impulsion électrique dans l'oreillette, et donc succession régulière et rapide de contractions sans pause.	La fréquence rapide environ 300bat/mn	l'oreillette et le ventricule ne sont pas synchronisés

Conclusion

Dans ce chapitre nous avons présenté d'une manière succincte les principales caractéristiques du signal cardiaque ECG normal et anormal ainsi que les intervalles et segments correspondants. Nous proposons dans le second chapitre une approche temporelle moyennant les réseaux de neurones RBF et l'algorithme GOFR (Generalized Orthogonal Forward Regression) pour modéliser chaque battement du signal cardiaque, par cinq gaussiennes avec des amplitudes, écarts types et moyennes différents. Ces derniers constitueront les paramètres pertinents d'un battement cardiaque qui seront utilisés par la suite pour classer les signaux cardiaques en normaux ou anormaux.

Introduction

Afin d'identifier automatiquement les ondes caractéristiques du signal cardiaque, on passe par une étape essentielle qui est la modélisation. L'objectif est donc de trouver un modèle mathématique optimal pour représenter un ECG, en essayant d'atteindre une erreur minimale entre le signal réel et le modèle.

La modélisation la plus couramment utilisée dans l'analyse de l'ECG est l'approximation du signal par une ligne brisée [38-40]. Les paramètres de ces lignes sont ensuite étudiés pour le repérage des QRS notamment, [41]. Cette représentation est loin de notre objectif, nous ne nous étendrons donc pas plus sur ce type d'analyse. D'autres méthodes de modélisation abondamment décrites dans la littérature reposent sur la décomposition du signal cardiaque par une transformée en ondelettes. On peut trouver des synthèses dans les travaux cités dans les références [42-43]. Il existe un grand nombre de types d'ondelettes telles que celles de Haar, de Morlet, de Daubechies, les « symlets » [44], etc. Le choix du type d'ondelettes dépend essentiellement des propriétés recherchées, par exemple la forme qui évoque celle des ondes cardiaques telle qu'une décomposition en ondelettes de type Coiflet (Mexican Hat) [44].

Une application de transformées en ondelettes pour l'analyse du signal ECG est proposée par [45] où le signal ECG est calculé par une application non linéaire qui met en évidence les points d'inflexions de la courbe correspondant à des débuts et des fins d'ondes potentielles. Une chaîne de Markov cachée analyse ensuite les résultats issus de cette décomposition pour repérer les ondes caractéristiques [46]. L'approche proposée qui consiste à associer à chaque régresseur (ici des ondelettes) une onde caractéristique n'est pas satisfaite par cette démarche. Une transformation du signal ECG en ondelette est également proposée par [47] pour la détection des ondes QRS, P et T. Une troisième application des ondelettes couramment proposée en ECG est la compression du signal [48-49], que nous exposerons dans le cadre de nos travaux en chapitre trois.

Pour notre part nous proposons une modélisation par réseau de neurone à fonction radiale de base « RBF » où la fonction d'activation est une gaussienne dont la forme peut correspondre aux ondes caractéristiques du signal ECG. Elle a comme paramètres la pondération A_i (amplitude), la moyenne μ_i et l'écart type σ_i qui sont les poids du réseau. Notre but est donc de concevoir un système d'aide au diagnostic, moyennant la modélisation mathématique des battements cardiaques par une approche temporelle utilisant les réseaux de neurones RBF (Radial Basis Function). Ce qui permettra l'extraction de paramètres temporels caractérisant chaque

battement du cœur .En effet , cerner ces ondes permet de fournir aux médecins des informations sur leurs formes, leurs distances mutuelles , de définir donc les seuils de normalité pour chacune des ondes et par conséquent de détecter les anomalies. Il faudra dès lors modéliser l'ECG en une somme de fonctions qui s'adaptent avec le plus de précision possible aux ondes caractéristiques. La modélisation telle que nous allons la présenter dans ce chapitre sera une opération incontournable dans n'importe quel processus lié à l'exploitation d'un signal cardiaque quelconque.

2.1 Modélisation par réseaux à fonctions de base radiales (RBF)

Introduit par Powell et Broomhead [50], le réseau RBF (Radial Basis Functions) fait partie des réseaux de neurones supervisés. Il est constitué de trois couches (figure 2.1): une couche d'entrée qui retransmet les entrées sans distorsion, une seule couche cachée qui contient les neurones RBF qui sont généralement des gaussiennes [50-52] et une couche de sortie dont les neurones sont généralement animés par une fonction d'activation linéaire. Chaque couche est complètement connectée à la suivante et il n'y a pas de connexions à l'intérieur d'une même couche. Les réseaux de RBF sont utilisés pour l'approximation de fonctions non linéaires. La fonction recherchée est toujours une combinaison linéaire de fonctions radiales. La propriété d'approximateurs universels pour ces réseaux n'a été que récemment prouvée pour des gaussiennes radiales et plus généralement pour des RBF. Ces réseaux seront donc utilisés comme outil de modélisation.

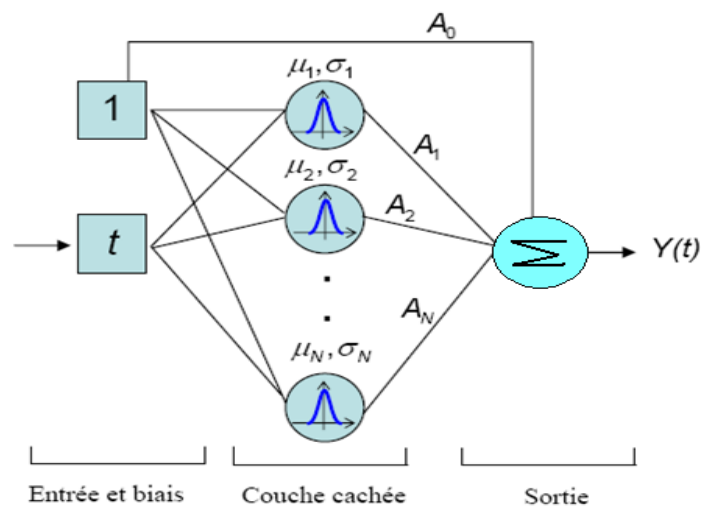


Figure 2.1: Réseau de neurones RBF avec une entrée temporelle.

2.1.1 Choix des RBF

Le choix des RBF repose sur le fait que toute fonction (continue sur un compact) peut être

approchée d'aussi près que l'on veut par une somme de fonctions gaussiennes judicieusement choisies, figure 2.2. Si l'on impose à ce que ces gaussiennes aient toutes la même hauteur, il suffit de remplacer "somme" par "combinaison linéaire" pour que la propriété soit conservée. La figure 2.2 représente une fonction $y = f(x)$, ainsi que les gaussiennes qui contribuent à sa reconstitution. L'objectif premier des réseaux RBF est de construire une bonne approximation d'une fonction qui n'est connue que par un nombre fini de points $\{x_i, y_i\}$.

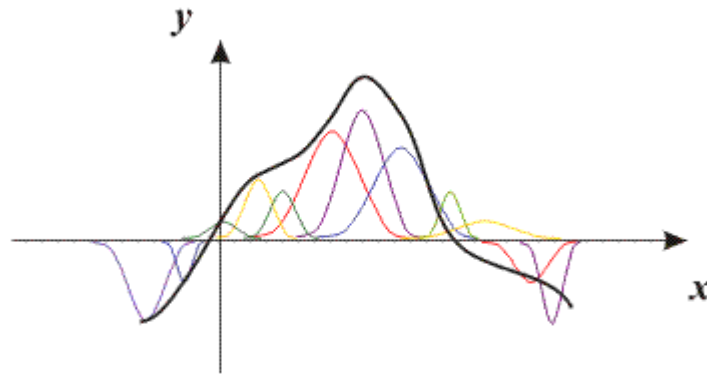


Figure 2.2 : Approximation d'une fonction par des gaussiennes.

2.1.2 Modélisation

La modélisation par RBF est pertinente pour l'étude qui nous intéresse car on peut choisir comme fonction de base une gaussienne qui, une fois ses paramètres adaptés, possède une forme qui pourrait être une bonne représentation de chacune des ondes cardiaques. Ici, les paramètres du réseau sont, d'une part, les pondérations des fonctions radiales, et, d'autre part, les paramètres des gaussiennes (moyenne et écart-type). Le nombre N de fonctions à utiliser pour le modèle doit être fixé a priori. En résumé, chaque gaussienne est représentée par les paramètres pertinents suivants : l'amplitude A_i , la moyenne μ_i et l'écart type σ_i qui sont les poids du réseau. Nous optimiserons ces paramètres afin de minimiser l'erreur quadratique moyenne appelée aussi le coût J .

2.1.2.1 Etapas de la modélisation

Notre modélisation qui a fait l'objet de plusieurs articles [54-56] consiste en :

- L'initialisation des paramètres des gaussiennes sélectionnées
- L'optimisation de ces paramètres

Dans la figure 2.1 ci-dessus de la modélisation par réseau de neurones RBF, la fonction de sortie sera de la forme suivante :

$$y(t) = \sum_{i=1}^N A_i \exp\left(-\frac{(t-\mu_i)^2}{2\sigma_i^2}\right) \quad (2.1)$$

Où : A_i est la pondération de l'ième gaussienne

μ_i est la moyenne de l'ième gaussienne

σ_i est l'écart type de l'ième gaussienne

N est le nombre de neurones ou de gaussiennes

Le modèle obtenu est la décomposition du signal ECG en une somme de fonctions gaussiennes.

Ce qui permettra d'associer à chaque onde caractéristique de l'ECG une fonction gaussienne.

Le coût J est défini par l'erreur quadratique moyenne:

$$J(\tilde{\mu}, \tilde{\sigma}, \tilde{A}) = \left(\frac{1}{N_e}\right) \sum_{i=1}^{N_e} (s(i) - y(i))^2 \quad (2.2)$$

$$J(\tilde{\mu}, \tilde{\sigma}, \tilde{A}) = \left(\frac{1}{N_e}\right) \sum_{i=1}^{N_e} \left(s(i) - \sum_{j=1}^N A_j \exp\left(-\frac{(j-\mu_j)^2}{2\sigma_j^2}\right)\right)^2 \quad (2.3)$$

Où $\tilde{\mu} = \{\mu_i\}_{i \in [1, N]}$, $\tilde{\sigma} = \{\sigma_i\}_{i \in [1, N]}$, $\tilde{A} = \{A_i\}_{i \in [1, N]}$

N_e : nombre d'échantillons du signal S à modéliser.

N : nombre de gaussiennes à sélectionner qui représente le nombre de neurones de notre réseau RBF.

La recherche des paramètres optimaux revient à minimiser le coût J qui est non linéaire. Pour ce faire, un algorithme itératif avec une initialisation arbitraire des paramètres du réseau a été utilisé, [53]. Ceci a conduit à multiplier le temps de calcul sans assurer la convergence du réseau vers un minimum global de J . Afin d'éviter cela, il est souhaitable que les valeurs initiales des paramètres soient estimées de telle façon à situer la fonction du coût J aussi proche que possible de son minimum.

2.1.2.2 Principe de l'algorithme GOFR (Generalized Orthogonal Forward Regression)

Le but de la modélisation de notre signal ECG par des gaussiennes est d'utiliser un seul régresseur pour représenter chaque onde de notre signal cardiaque. Pour cela, nous avons fait appel à une méthode de sélection adaptative appelée GOFR, [54,56,57]. L'utilisation de cet algorithme nécessite la construction d'une bibliothèque B qui contient des gaussiennes de différentes largeurs σ_i centrées à différents emplacements temporels μ_i .

Dans l'algorithme OFR (Orthogonal Forward Regression) fondé sur l'algorithme d'orthogonalisation de Gram-Schmidt, les régresseurs pertinents sont d'abord sélectionnés, puis ajustés au battement, [58-59]. La procédure GOFR que nous avons appliquée, intègre à chaque étape sélection et ajustement, régresseur par régresseur et avant chaque orthogonalisation par rapport à un régresseur sélectionné, une optimisation sommaire des paramètres de celui-ci afin de le rapprocher au mieux du signal. Une fois optimisé, ce nouveau régresseur remplace l'ancien dans la bibliothèque et l'orthogonalisation se fait par rapport à ce dernier.

Les différentes étapes de l'algorithme GOFR se résument à :

- Construction de la bibliothèque
- Sélection de la Gaussienne la plus pertinente B_i
- Optimisation des paramètres de la Gaussienne
- Orthogonalisation du battement
- Orthogonalisation de la bibliothèque

Le battement à modéliser S et les régresseurs de la bibliothèque sont des signaux temporels constitués de N_e points d'échantillonnage. Pour l'application de l'algorithme GOFR, ils sont considérés comme des vecteurs de l'espace à N_e dimensions, appelé habituellement « espace des observations ». On considérera donc deux représentations :

- ❖ une représentation temporelle, dans laquelle le signal S et les régresseurs sont des fonctions du temps, cet espace est appelé espace des variables,
- ❖ et une représentation vectorielle, dans laquelle le signal S et les régresseurs B sont des vecteurs de l'espace des observations, de dimension N_e .

Soit Y le modèle construit, et E l'erreur de modélisation. Avant la première itération, le modèle construit est $Y = 0$; l'erreur de modélisation est donc $E^1 = S - 0 = S$. Dans la représentation temporelle, le signal, noté S , est constitué des N_e points d'échantillonnage, et les régresseurs sont des fonctions temporelles paramétrées. L'apprentissage, qui porte sur les paramètres de ces régresseurs, est donc réalisé pour cette représentation, et c'est également dans cette représentation que se calcule l'erreur quadratique moyenne J entre le signal S et le modèle Y , et l'erreur de modélisation E .

Dans la représentation vectorielle, le signal S est un vecteur de N_e dimensions, ainsi que chacun des régresseurs. On effectue le calcul des angles pour la sélection itérative des N régresseurs du modèle. Le numéro de l'itération $i = 1..N$ est indiqué par un exposant. On ajoute

une (*) à droite du symbole des régresseurs indiquant que ceux-ci ont été ajustés par optimisation.

Ainsi le signal qui permet la sélection du régresseur à la i -ème itération dans l'espace des observations : S^i , n'est plus relié à l'erreur de modélisation E^i qui permet l'apprentissage de cette gaussienne une fois sélectionnée. L'hypothèse faite ici est que le S^i de la représentation vectorielle est proche de la partie de E^i qui est dans l'espace orthogonal aux régresseurs d'indices (m_1, m_{i-1}) .

2.1.2.3 Construction de la bibliothèque de gaussiennes B

Pour un signal de N_e points, B est l'ensemble suivant :

$$B = \left\{ \begin{array}{l} G(\mu, \sigma) \\ \mu = j * \frac{N_e}{2^k} \\ \sigma = \frac{N_e}{2^{k+1}} \\ j \in [0, 2^k] \\ k \in [1 \dots K_{\max}] \end{array} \right\} \quad (2.4)$$

- $G(\mu, \sigma)$ est la fonction gaussienne d'écart-type σ et de moyenne μ .

$$G(\mu, \sigma) = e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

- $K_{\max} = 6$.
- $N_e = 342$, nombre d'échantillons du signal.
- $N_b = 132$ où N_b est le nombre de régresseurs.

Dans la suite des équations, les N_b régresseurs ou gaussiennes de la bibliothèque B seront notés $\{B^i\}_{i=1, N_b}$, [56].

Chaque itération de notre algorithme est constituée de trois étapes.

➤ Première itération

Les trois étapes sont :

- **la sélection** du meilleur régresseur B_{m_1} de la bibliothèque obtenue par le calcul des angles que chacun d'eux forme avec le signal.
- **L'ajustement** des paramètres du régresseur sélectionné pour obtenir le nouveau régresseur $B^*_{m_1}$.
- **L'orthogonalisation** des régresseurs restants et du signal par rapport à $B^*_{m_1}$ pour se placer dans un espace orthogonal à ce dernier noté $\perp B^*_{m_1}$. Les itérations suivantes sont construites sur ce même schéma de trois étapes.

🚦 Selection

À la première itération, le signal est $S = S^1$, et les vecteurs régresseurs sont les $B^1 = \{B^1_i\}_{i=[1..Nb]}$; le vecteur de la famille B^1 le plus pertinent c'est celui qui fait le plus petit angle avec le vecteur signal. Autrement dit, qui a le plus grand cosinus en valeur absolue avec le signal S^1 . Il porte l'indice m_1 et s'écrit de la manière suivante :

$$m_1 = \operatorname{argmax} \left(\frac{|\langle S^1 | B^1_k \rangle|}{\|S^1\| \cdot \|B^1_k\|} \right) \quad (2.5)$$

Une fois sélectionné, le régresseur B_{m_1} peut être envisagé comme une fonction temporelle paramétrée utilisée pour modéliser S ; on optimise ses paramètres pour le rapprocher au mieux du signal compris dans une fenêtre temporelle dont les paramètres seraient μ_{m_1} , σ_{m_1} avant de chercher un deuxième régresseur.

🚦 Optimisation du régresseur 1

L'optimisation du régresseur $B^1_{m_1}$ se fait dans la représentation temporelle en minimisant l'erreur quadratique moyenne J entre ce qui est à modéliser: $E^1=S$, et la fonction B_{m_1} d'écart-type σ_{m_1} , de position temporelle μ_{m_1} , et d'amplitude A_{m_1} .

$$J(\sigma_{m_1}, \mu_{m_1}, A_{m_1}) = \left(\frac{1}{Ne} \sum_{k=1}^{Ne} \left(E^1(k) - A_{m_1} \exp \left(-\frac{(k - \mu_{m_1,1})^2}{2\sigma_{m_1}^2} \right) \right) \right)^2 \quad (2.6)$$

Cette optimisation a l'avantage d'être rapide car on n'optimise ici que trois paramètres: les paramètres d'une gaussienne.

A l'issue de cette optimisation, les paramètres $\sigma^*_{m_1}$, $\mu^*_{m_1}$ et $A^*_{m_1}$ du régresseur $B^*_{m_1}$ sont désormais ajustés, et sont solutions du problème d'optimisation.

🚦 Orthogonalisation

L'orthogonalisation (figure 2.3) est effectuée par rapport au régresseur ajusté $B^{1*}_{m_1}$ et non par rapport au régresseur sélectionné $B^1_{m_1}$.

$$B_k^{i+1} = B_k^i - \frac{\langle B_k^i | B_{mi}^i \rangle}{\langle B_{mi}^i | B_{mi}^i \rangle} B_{mi}^i \quad \forall k \in [1, N_b] \dots \quad (2.7)$$

$$S^{i+1} = S^i - \frac{\langle S^i | B_{mi}^i \rangle}{\langle B_{mi}^i | B_{mi}^i \rangle} B_{mi}^i \dots \quad (2.8)$$

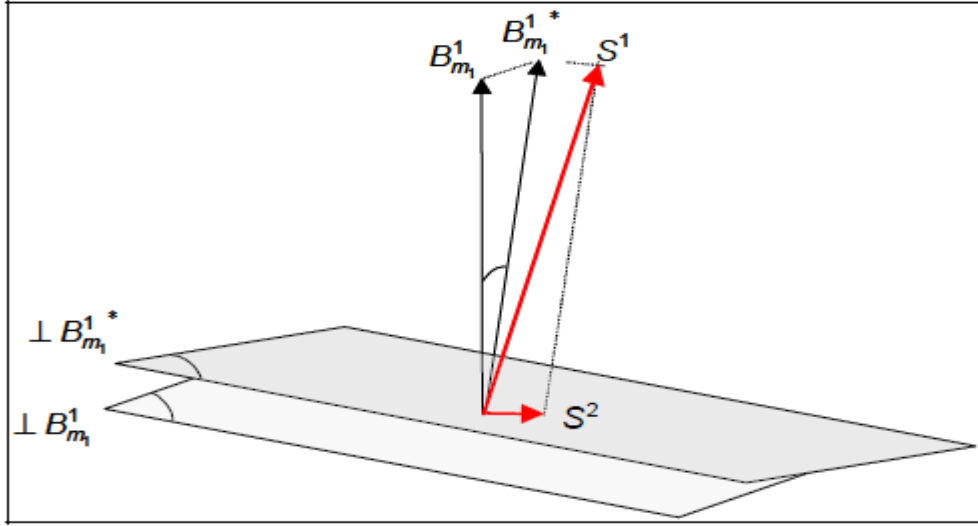


Figure 2.3 : Orthogonalisation par GOF.

S^{i+1} et les $\{B_k^{i+1}\}_{k=[1..N_b] \setminus \{m1\}}$ sont respectivement ce qui reste du signal et des régresseurs dans L'espace orthogonal à B_{m1}^* .

S^2 et les $\{B_k^2\}_{k=[1, N_b] \setminus \{m1\}}$ sont respectivement ce qui reste du signal et des régresseurs de leurs projection dans l'espace orthogonal à B_{m1}^* .

$$B_k^2 = B_k^1 - \frac{B_k^1 * B_{m1}^{1*}}{B_{m1}^{1*} * B_{m1}^{1*}} * B_{m1}^{1*} \quad \forall k \in [1, N_b] - \{m1\} \quad (2.9)$$

$$S^2 = S^1 - \frac{S^1 * B_{m1}^{1*}}{B_{m1}^{1*} * B_{m1}^{1*}} * B_{m1}^{1*} \quad (2.10)$$

En revenant à la représentation temporelle, l'erreur de modélisation E_2 à l'itération 2 s'écrit :

$$E_2(t) = S(t) - Y(t) = S(t) - B_{m1}^*(t) \quad (2.11)$$

Le modèle étant à cette étape :

$$Y(t) = B_{m1}^*(t) \quad (2.12)$$

Après la première orthogonalisation, l'algorithme opère dans un espace de dimension N_e-1 , et nous disposons de N_b-1 régresseurs pour traiter le signal S_2 et donc préparer l'itération suivante (sélection, optimisation, orthogonalisation).

Les itérations suivantes sont construites sur ce même schéma de trois étapes :

- La sélection du meilleur régresseur B_{mi} de la bibliothèque obtenue par le calcul des angles que chacun d'eux forme avec le signal.
- L'ajustement des paramètres du régresseur sélectionné pour obtenir le nouveau régresseur B^*_{mi} .
- L'orthogonalisation des régresseurs restants et du signal par rapport à B^*_{mi} .

Toutes ces étapes d'orthogonalisation se feront, après optimisation de chaque régresseur moyennant l'algorithme d'optimisation du premier ordre, autant de fois qu'il y a d'ondes à modéliser.

En effectuant N itérations de cet algorithme, on obtient N gaussiennes pour modéliser le signal S . Le modèle Y ainsi obtenu peut s'écrire :

$$Y(t) = \sum_{i=1}^N B^*_{mi}(t) \quad (2.13)$$

où les $\{B^*_{mi}\}_{i=1, N_b}$ sont les gaussiennes de paramètres $\{(\sigma^*_{mi}, \mu^*_{mi}, A^*_{mi})\}$.

2.2 Résultats de la modélisation

Notre réseau de neurones est conçu de façon à modéliser n'importe quel signal ECG. Les signaux utilisés comptent une centaine d'enregistrements à deux dérivations (deux voies), extraits des bases de données Physio Bank, développée en grande partie par l'institut de technologie du Massachusetts (MIT) et l'hôpital Beth Israël de Boston (BIH) [60]. Dans notre cas nous avons exploité la base de données ECG Databases. Elle est constituée d'un ensemble d'enregistrements incluant un ou plusieurs signaux digitalisés d'ECG et un ensemble d'annotations de battements, échantillonnés à 360 Hz [60]. C'est une base de données utilisée par les chercheurs, qui permettra de comparer nos résultats et ainsi leur donner une bonne interprétation.

Pendant l'enregistrement des signaux électrocardiographiques, des bruits de sources diverses peuvent être superposés au signal d'origine. Le prétraitement des signaux ECG s'impose pour l'extraction de ce signal.

A sa sortie de l'électrocardiographie, le signal a déjà subi deux filtrages:

- Un filtre passe haut, pour éliminer les bruits de haute fréquence.

- Un filtre réjecteur, pour se débarrasser de la fréquence du secteur "50hz".

2.2.1. Prétraitement

Un prétraitement numérique des signaux ECG de la base de données MIT Data Base ECG, a été réalisé [61]. Ce traitement a nécessité l'application des étapes suivantes :

- L'analyse de la ligne de base en utilisant le filtrage dans les basses fréquences
- Détection du complexe QRS moyennant des filtres numériques
- Le découpage qui se fait essentiellement par la détection d'ondes

La première phase de prétraitement consiste à éliminer les nuisances introduites par la ligne de base qui est une importante variation de la ligne isoélectrique qui apparaît généralement lorsque le patient n'effectue aucun mouvement et que le cœur n'a aucune activité électrique. Pour l'analyse d'un enregistrement ECG, et dans l'objectif d'un traitement automatique d'un tel signal, il est impératif de la repérer précisément pour fixer le « zéro ». Cela permet une comparaison simple des amplitudes des ondes d'un battement à l'autre.

Pour chaque enregistrement de 10s (figure 2.4), nous avons éliminé d'abord la ligne de base puis procédé au découpage des signaux ECG en battements. La méthode consiste à filtrer notre signal ECG d'origine avec un filtre passe bas de fréquence de coupure de 2hz. Suite à cette première étape, nous réalisons la soustraction du signal d'origine avec son signal filtré afin d'en extraire la ligne de base puis de la soustraire au signal d'origine. Ces deux étapes sont équivalentes à un filtrage passe haut (figure 2.5) [62].

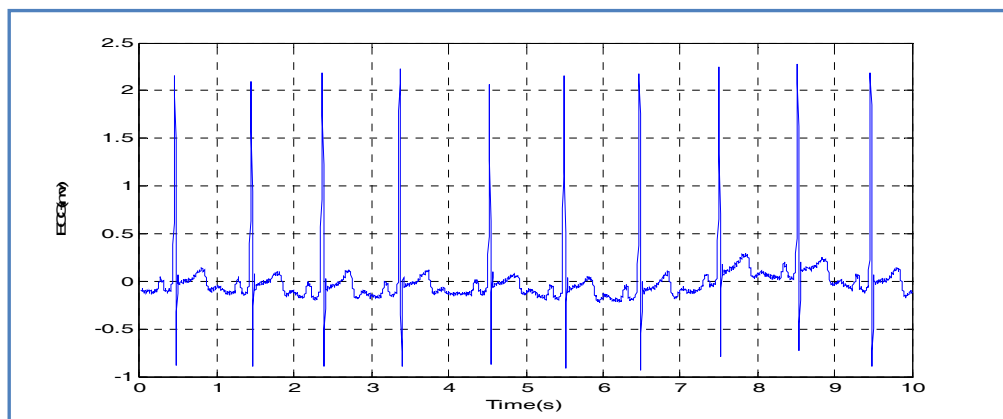


Figure 2.4: Signal ECG avec ligne de base.

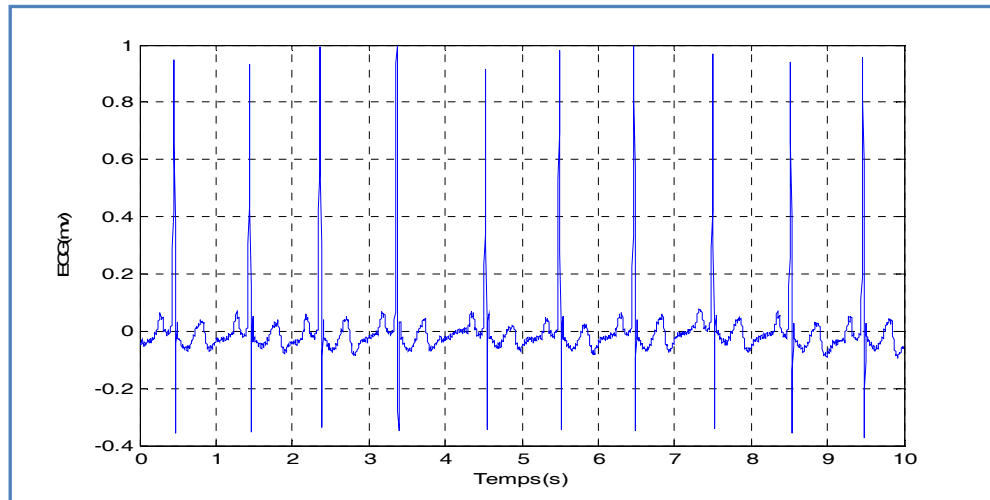


Figure 2.5 : ECG1.dat corrigé de sa ligne de base.

Les résultats obtenus nous permettent dans la phase suivante de détecter les pics R. Il est à remarquer que ce filtrage est nécessaire pour la détection du complexe QRS.

Pour cela, nous avons commencé par appliquer l'algorithme de Pan et Tompkins [63], qui consiste en une série de filtres numériques (figure 2. 6).

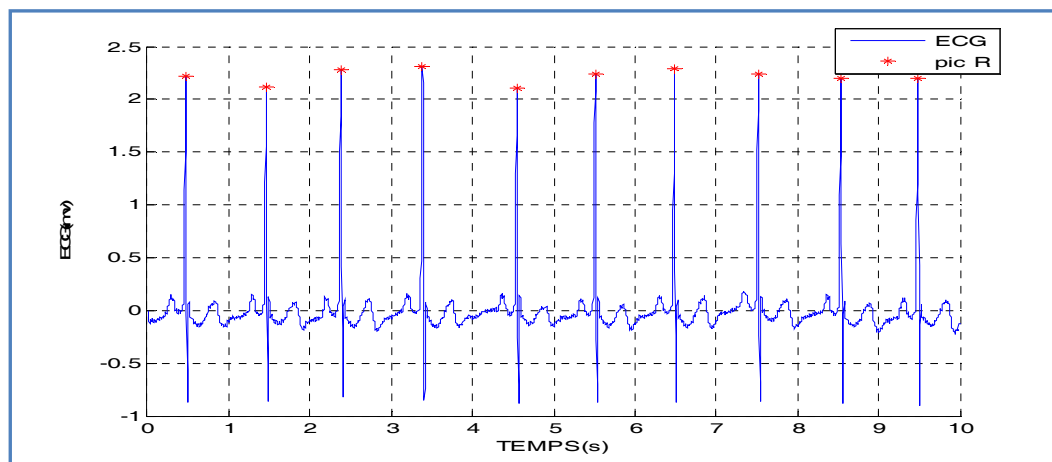


Figure 2.6 : Détection des QRS (détection des pics R).

Après extraction de l'intervalle RR lors de la phase précédente, nous avons détecté l'onde T (figure2.7). Pour effectuer cette détection, nous avons défini une fenêtre à partir de la position du pic de l'onde R (pos R). La taille de cette fenêtre est fonction de la durée du segment RR.

Les différentes étapes de détection de l'onde T sont:

- Lire le segment RR.
- Définir la taille de la fenêtre de l'onde T:

1. Si $RR > 1300$ ms:

$$T.F = 0.1833 * RR + 175$$

2. Si $RR < 600$ ms:

$$T.F = 0.208 * RR + 158$$

3. Dans le reste des cas:

$$T.F = 11.5 * \sqrt{RR}$$

Début de la fenêtre

$$INI = R + T.F/2.5$$

Fin de la fenêtre:

$$FIN = INI + T.F$$

- $T = ECG * \text{fenêtre}$

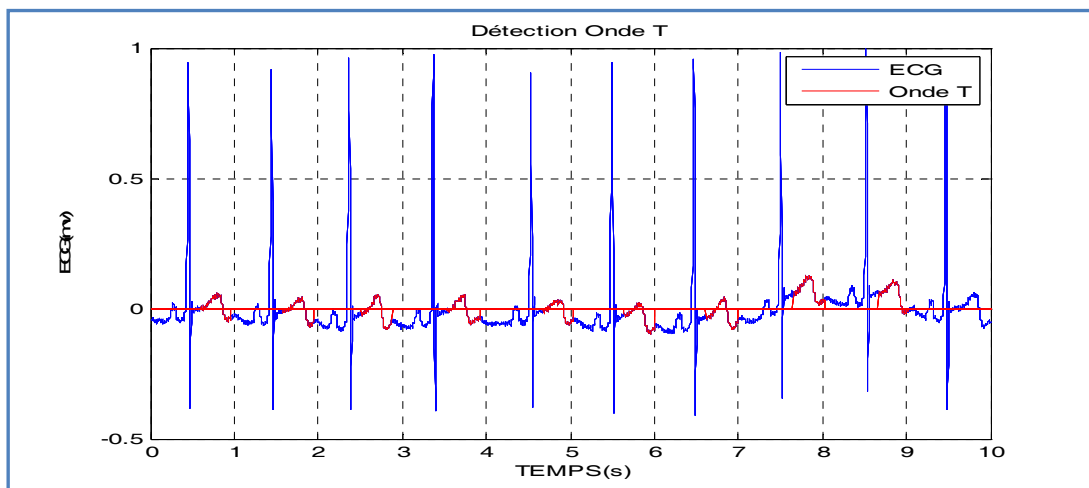


Figure 2.7 : Détection de l'onde T.

A partir de la détection de l'onde T, nous avons effectué le découpage des signaux ECG en plusieurs battements. Chaque battement s'étend de la fin de l'onde T à la fin de l'onde T du battement suivant (figure 2.8).

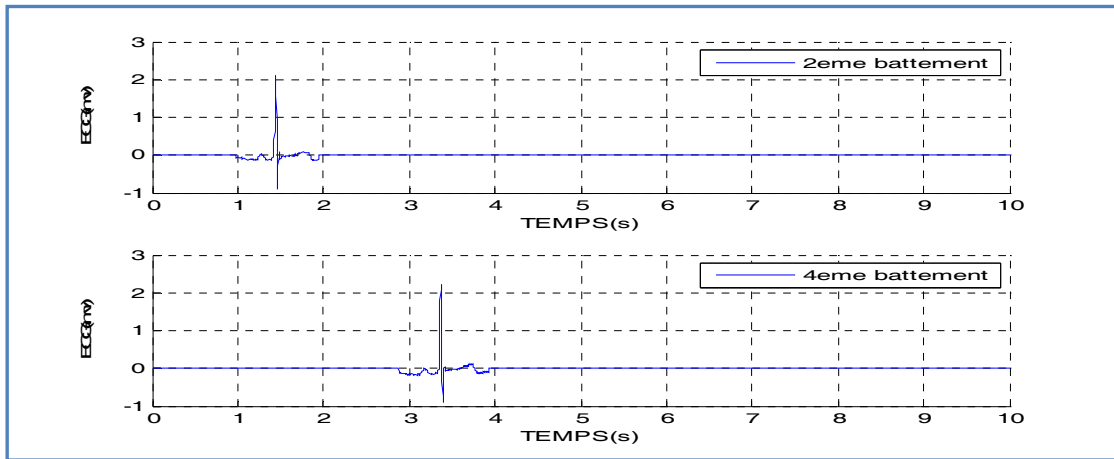


Figure 2.8: ECG après découpage.

Pour un signal ECG donné, chaque battement détecté sera modélisé selon l'algorithme GOFR.

2.2.2 Résultats graphiques de la bibliothèque de gaussiennes

Dans la bibliothèque B de l'équation (2.4), nous avons fixé $K_{max} = 6$ pour avoir une riche bibliothèque avec $N_b=132$ gaussiennes, [56].

$B^i(\mu_i, \sigma_i)$ est la fonction gaussienne d'écart-type (σ_i) et de moyenne (μ_i). Chaque graphe représente la famille de gaussiennes qui ont une même largeur et qui balayent les 342 points d'échantillonnage du signal.

Lors de nos travaux, des tests ont été effectués pour estimer le nombre N_e de façon optimale. En variant le nombre de points, le choix a été fait en fonction d'une erreur quadratique moyenne minimale. Les résultats obtenus ont prouvé que l'erreur quadratique moyenne augmente pour $N_e \geq 343$. Le nombre d'échantillons choisi a été donc de 342 points. Ceci sera illustré ultérieurement.

Chaque graphe de la figure 2.9 à la figure 2.14, représente la famille de gaussiennes qui ont une même largeur et qui balayent les 342 points du signal. La bibliothèque établie précédemment pour $k=1$ jusqu'à $k=k_{max}$ fixé à 6 est représentée ci-dessous.

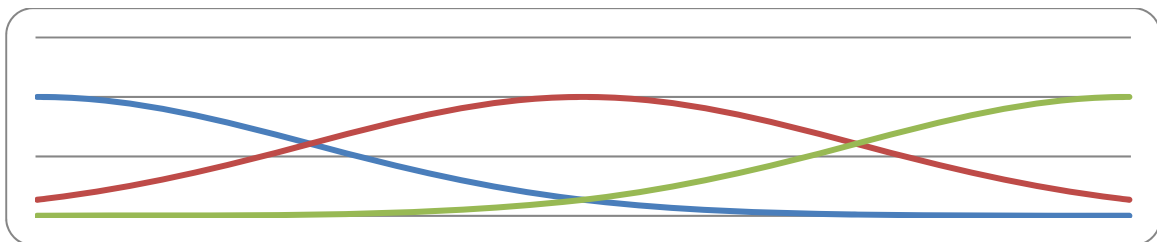


Figure 2.9: les gaussiennes de la bibliothèque pour $k = 1$: L'écart type $\sigma=85.5$.

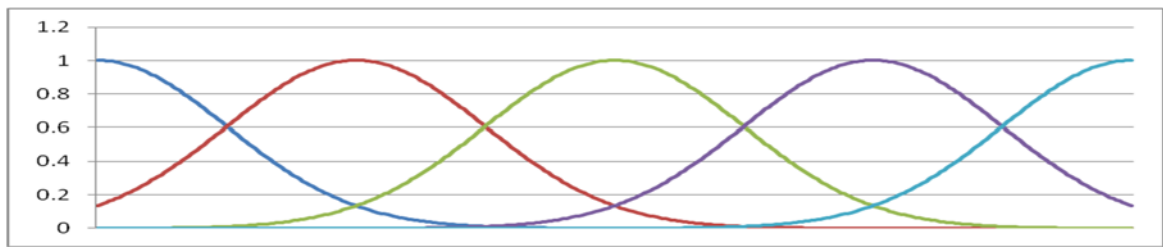


Figure 2.10: les gaussiennes de la bibliothèque pour $k = 2$: L'écart type $\sigma=42.75$.

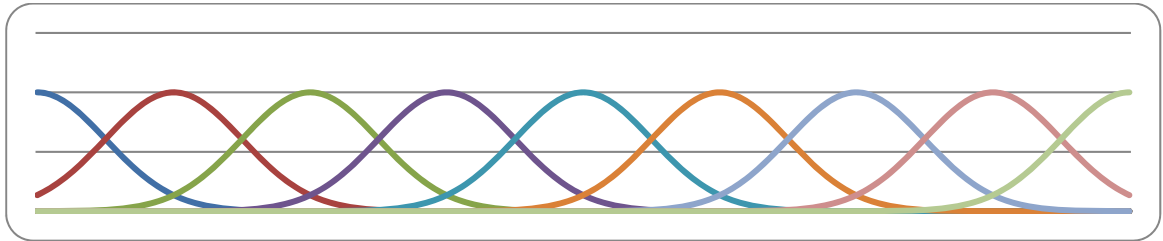


Figure 2.11: les gaussiennes de la bibliothèque pour $k = 3$: L'écart type $\sigma=21.375$.

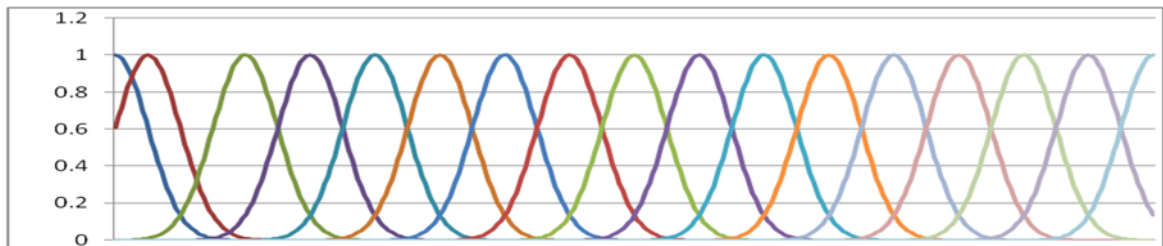


Figure 2.12: les gaussiennes de la bibliothèque pour $k = 4$: L'écart type $\sigma=10.6875$.

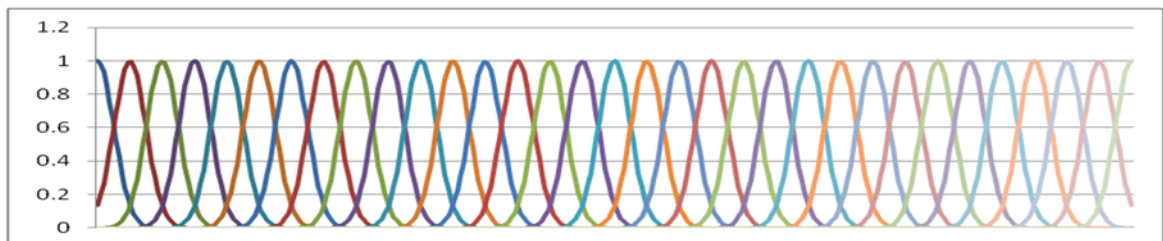


Figure 2.13: les gaussiennes de la bibliothèque pour $k = 5$: L'écart type $\sigma=5.34375$.

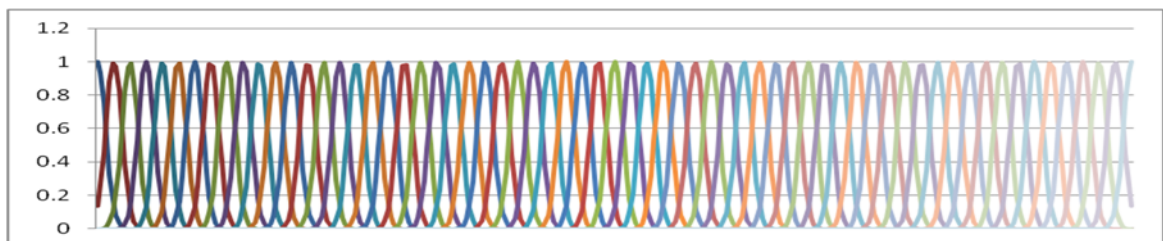


Figure 2.14: les gaussiennes de la bibliothèque pour $k = 6$: L'écart type $\sigma=2.671875$.

Les résultats de la bibliothèque permettraient de sélectionner exactement les gaussiennes les plus pertinentes afin de modéliser un signal ECG donné.

Pour les ondes caractéristiques du signal de petite largeur, nous avons plus de choix que pour les ondes de grande largeur.

2.2.3. Sélection de la Gaussienne

Le choix de la gaussienne à partir de notre bibliothèque déjà construite, est basé sur le critère de ressemblance. Nous essayons de trouver la gaussienne qui pourrait représenter au mieux notre onde caractéristique, et ceci grâce à l'équation (2.5) décrite précédemment.

Dans notre cas, la gaussienne choisie (Figure 2.15) porte le numéro 53.

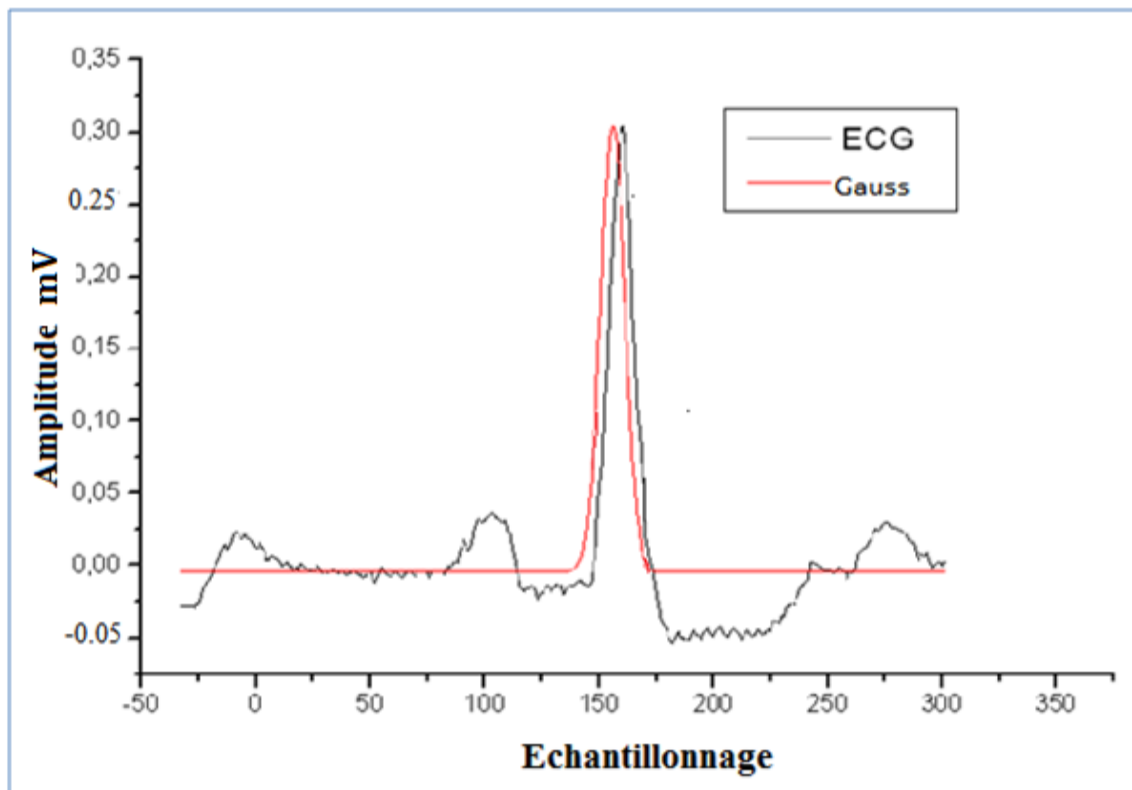


Figure 2.15 : Première gaussienne sélectionnée.

Nous remarquons que notre gaussienne sélectionnée a une moyenne très proche de l'onde que nous voulons modéliser, ce qui permet de ne pas avoir de conflit lors de l'optimisation

2.2.4. Optimisation paramètre par paramètre

L'optimisation est réalisée par l'algorithme du Gradient Projeté [54]. C'est un algorithme de

minimisation du premier ordre appliqué aux problèmes d'optimisation sous contraintes [64]. Cette optimisation prend en compte seulement trois paramètres: A_{m1} , σ_{m1} , μ_{m1} de la Gaussienne sélectionnée.

Une fois l'optimisation réalisée, les paramètres σ_{m1}^* , μ_{m1}^* et A_{m1}^* de la Gaussienne optimale B_{m1}^* sont ajustés, et représentent les solutions du problème d'optimisation.

➤ Ajustement de l'amplitude

A partir de notre algorithme, l'ajustement de l'amplitude se fait par opération mathématique de multiplication de notre vecteur (gaussienne choisie), par l'amplitude max de notre onde réelle détectée.

L'amplitude de notre gaussienne sélectionnée à partir de la bibliothèque étant de 1, sa multiplication par le maximum de l'onde détectée, met les deux ondes (gaussienne et onde réelle détectée) au même niveau comme c'est le cas sur la Figure 2.15 ci-dessus.

➤ Ajustement de la moyenne

En se basant sur le cas le plus simple d'une onde (onde ayant un pic représenté par un seul point et non-pas un segment), on fait une translation de notre onde sur l'axe des abscisses jusqu'à ce que ce max corresponde au max de l'onde sélectionnée de notre ECG (figure 2.16).

➤ Ajustement de l'écart type

Une fois l'amplitude et la moyenne ajustées, on passe à l'ajustement de notre écart type. Cet ajustement se fait grâce à l'erreur quadratique J de l'équation (2.6). On élargit notre régresseur autant de fois qu'il faudra jusqu'à obtenir l'erreur quadratique minimale.

A l'étape d'optimisation, une seule onde caractéristique du battement (P,Q, R, S or T) est modélisée par la Gaussienne B_{m1}^* . Le modèle à la première itération est:

$$y(k) = A_i^* \exp\left(-\frac{(k - \mu_i^*)^2}{2\sigma_i^{*2}}\right) \quad (2.14)$$

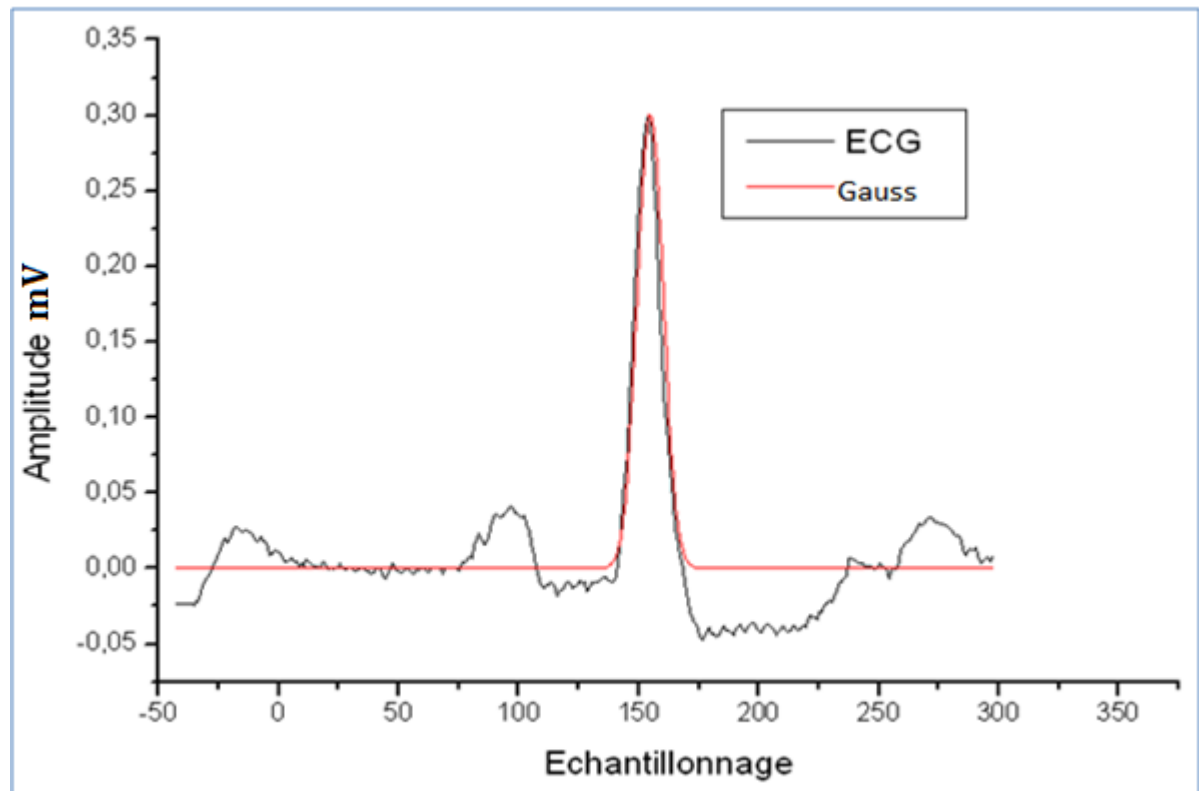


Figure 2.16 : Première onde après optimisation.

2.2.5 Orthogonalisation

Nous réalisons deux orthogonalisations: une, pour le battement original S et la seconde pour toutes les Gaussiennes de la bibliothèque B . L'orthogonalisation est effectuée par rapport au régresseur ajusté B_{m1}^1 . Après la première orthogonalisation, nous avons $N_b - 1$ Gaussiennes pour modéliser le signal S^2 , qui est utilisé pour les itérations suivantes (sélection, optimisation et orthogonalisation).

Les itérations suivantes sont construites sur ce même schéma de trois étapes :

- La sélection du meilleur régresseur B_{mi} de la bibliothèque.
- Ajustement de tous les paramètres du régresseur sélectionné pour obtenir la nouvelle gaussienne B_m^* .
- L'orthogonalisation des régresseurs restants et du signal par rapport à B_{mi}^* (optimisation).

Chaque onde cardiaque caractéristique P, Q, R, S et T concorde exactement avec une fonction Gaussienne. A chaque itération, l'algorithme modélise une onde du battement seulement.

Après cinq itérations de chacune des quatre étapes de l'algorithme, le battement a été modélisé par des Gaussiennes comme le montre la figure 2.17.

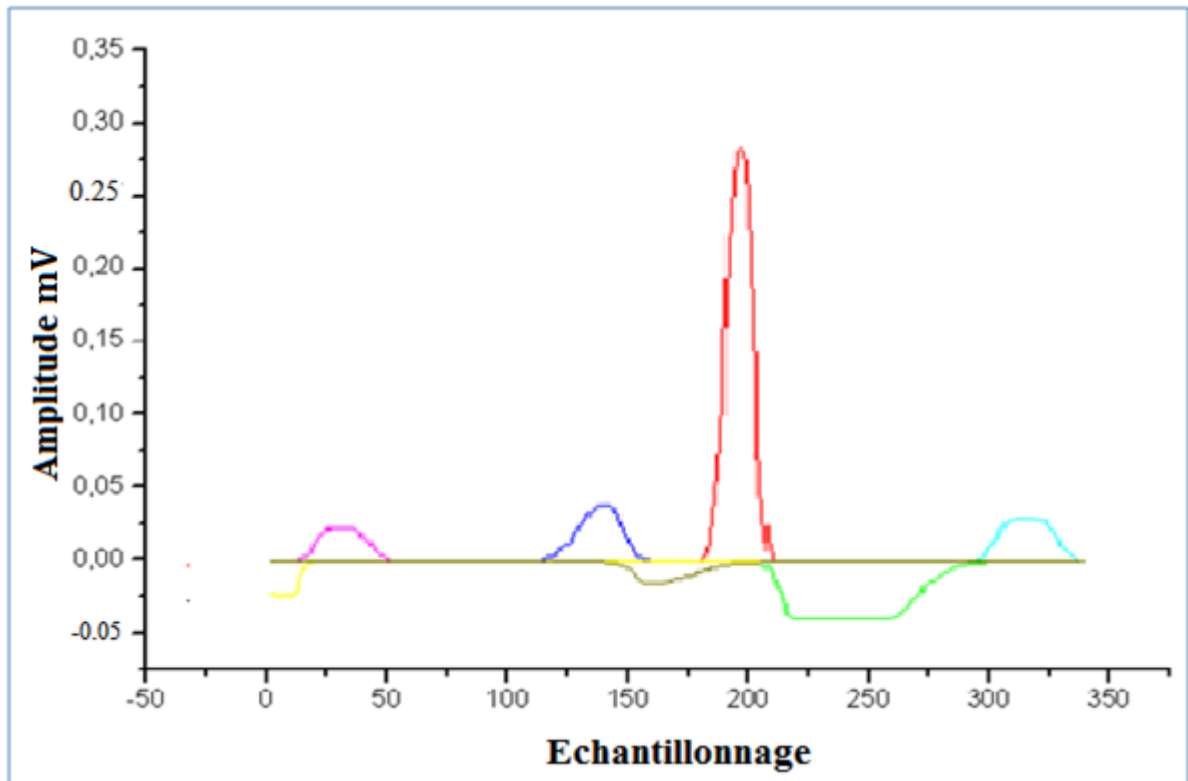


Figure 2.17: Modèle du battement cardiaque.

Finalement, ce modèle est une combinaison linéaire des cinq fonctions Gaussiennes qui représentent chaque onde caractéristique du battement cardiaque [56].

Après addition des ondes modélisées, on remarque que notre signal modélisé correspond parfaitement à l'exception de quelques points au signal réel (figure 2.18), avec une erreur quadratique globale J de l'ordre de 4.10^{-5} .

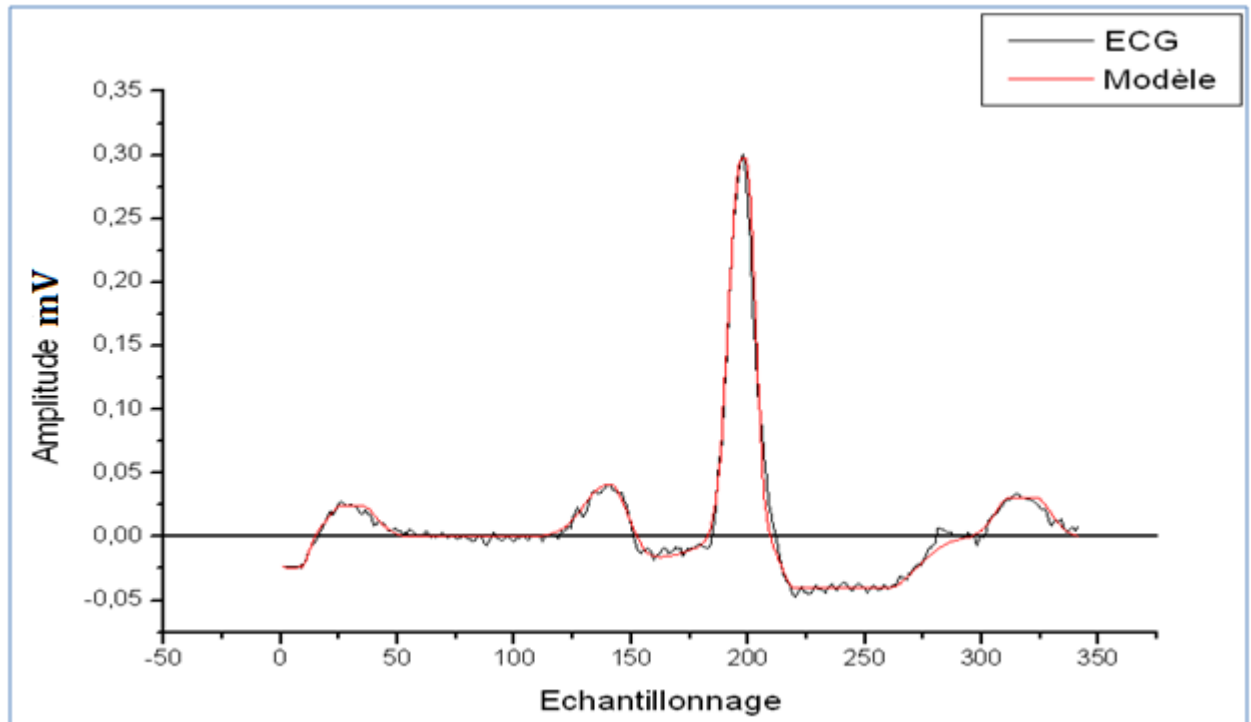


Figure 2.18 : Tracés du signal ECG réel et du modèle.

Afin de valider nos résultats de modélisation, nous avons appliqué les étapes de modélisation sur plusieurs signaux. A titre d'exemple, nous prenons le signal cardiaque ECG2 de la figure2.19.

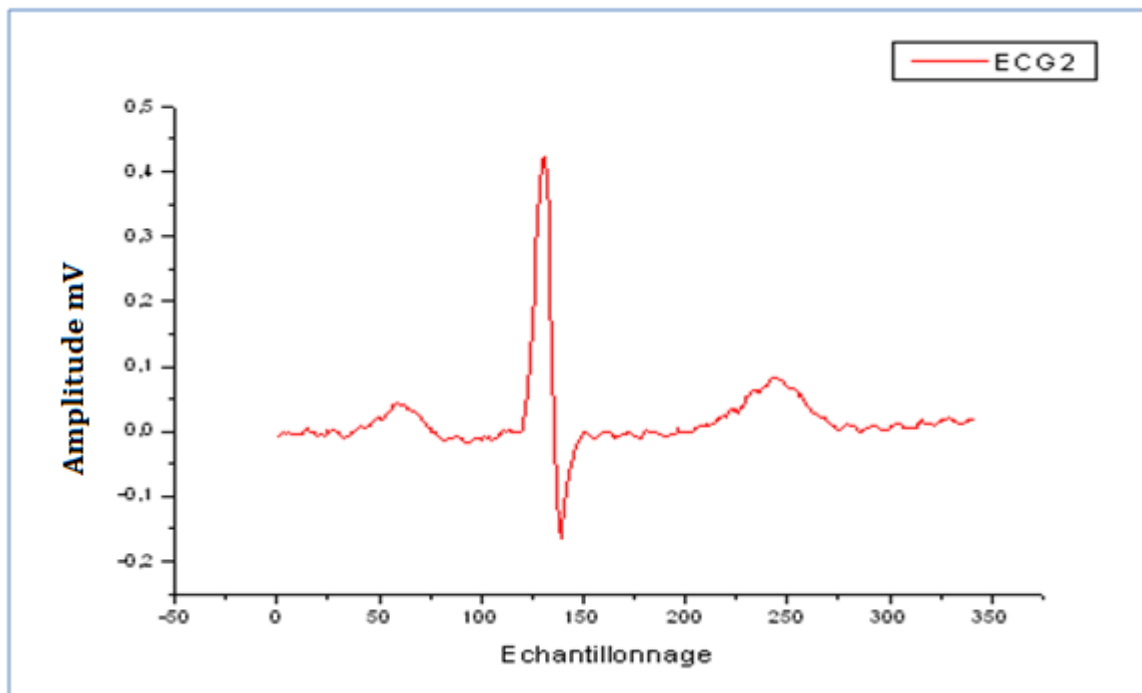


Figure 2.19 : Tracé du signal ECG2 réel.

Les résultats sont illustrés de la même manière que précédemment.

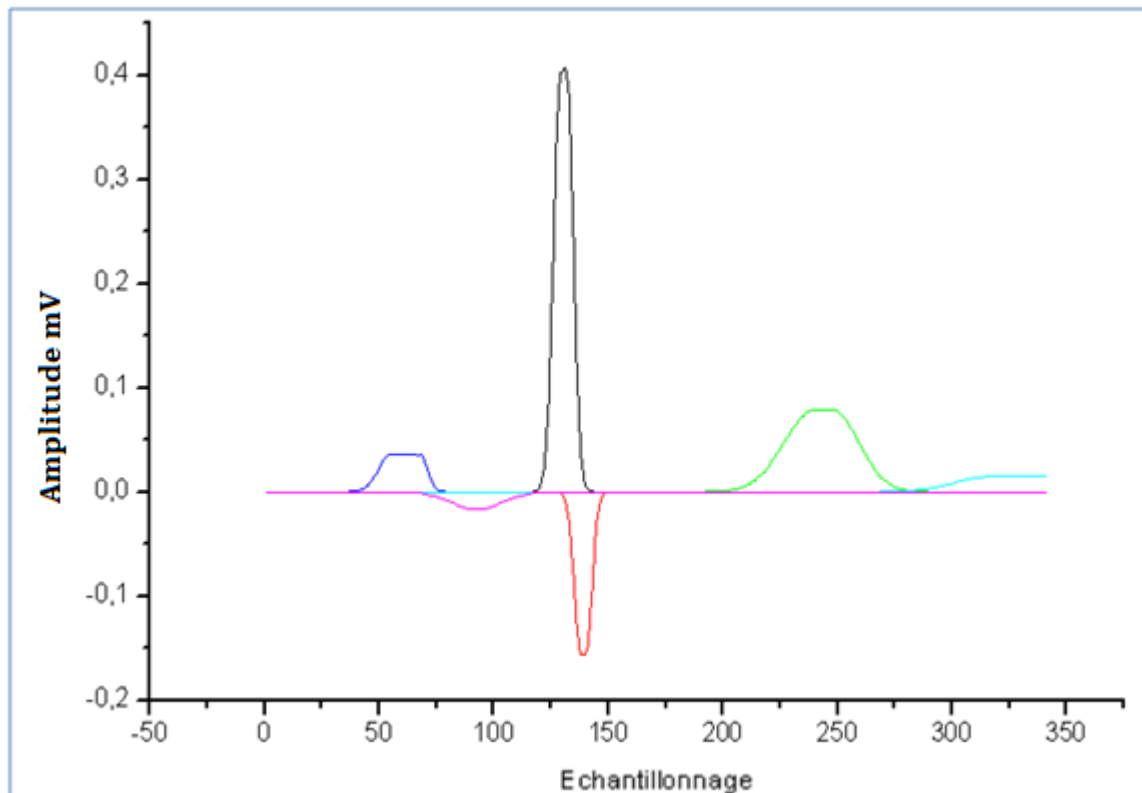


Figure 2.20: Modèle de l'ECG2.

La figure 2.20 qui représente le modèle de l'ECG2, découle des différentes étapes de modélisation : sélection, optimisation et orthogonalisation. Après addition des ondes modélisées, dans la figure 2.21 on établit la superposition du modèle avec l'ECG2 réel. On remarque que notre signal modélisé est obtenu, avec une erreur quadratique globale J de même ordre que précédemment, soit 7.10^{-5} . Le modèle 2 se rapproche étroitement de l'ECG2 réel.

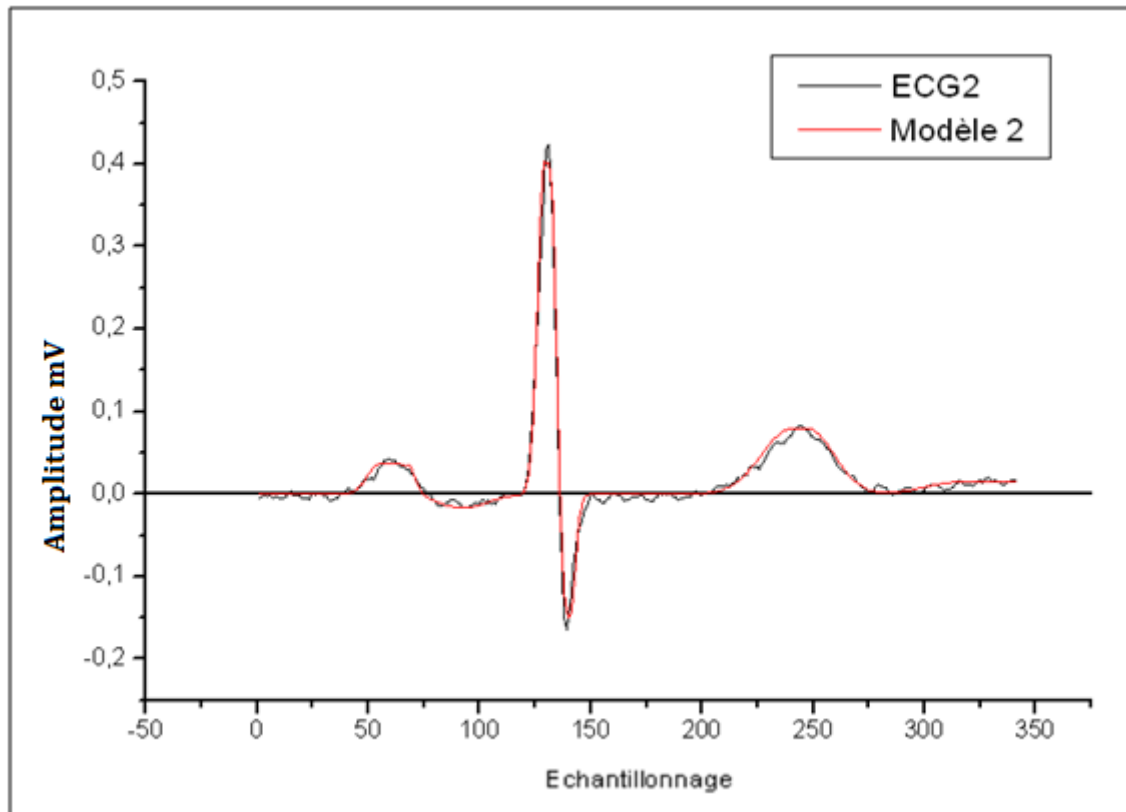


Figure 2.21 : Résultat final de la modélisation de l'ECG2.

Grâce à l'algorithme GOFr, les gaussiennes prédéfinies dans notre bibliothèque sont classées par ordre de pertinence, par rapport au signal cardiaque à modéliser. Le réseau de neurones est initialisé par les paramètres des N gaussiennes les plus pertinentes. Adapter les paramètres des gaussiennes sélectionnées par GOFr au tronçon de l'électrocardiogramme à modéliser, revient à optimiser les paramètres $\{A_i\}$, $\{\mu_i\}$ et $\{\sigma_i\}$ et donc minimiser le coût J qui est non linéaire en ces paramètres. Les résultats obtenus ont permis de réaliser une modélisation du signal ECG avec une erreur quadratique globale J de l'ordre de 10^{-5} en garantissant une reproduction approuvée par les spécialistes et donc cliniquement satisfaisante.

2.3 Résultats justifiant le choix du nombre de points d'échantillonnage N_e

Tout d'abord, nous rappelons que N_e est le nombre de points par battement du signal à modéliser. Nous rappelons aussi que ce choix est très important lors de la phase de découpage. Il n'y a aucun critère et aucune loi précise qui nous guide pour estimer le nombre N_e de façon optimale. C'est pour cette raison que nous avons effectué plusieurs tests afin de faire un choix

judicieux ; c'est-à-dire pour lequel l'erreur quadratique moyenne est minimale. A titre d'illustration pour notre thèse, nous avons utilisé les signaux :

ECG 6 : [ARRITHMIA-db 100 MIT-DB].

ECG 105: [ARRITHMIA-db 105 MIT-DB].

Sur lesquels nous avons fait des tests en variant N_e et en calculant à chaque fois l'erreur quadratique moyenne de modélisation sur chaque signal une fois reconstitué. Les résultats sont donnés sous forme graphique sur les figures 2.22 et 2.23. Les appellations ; tronçon ou battement d'un signal ECG sont équivalentes.

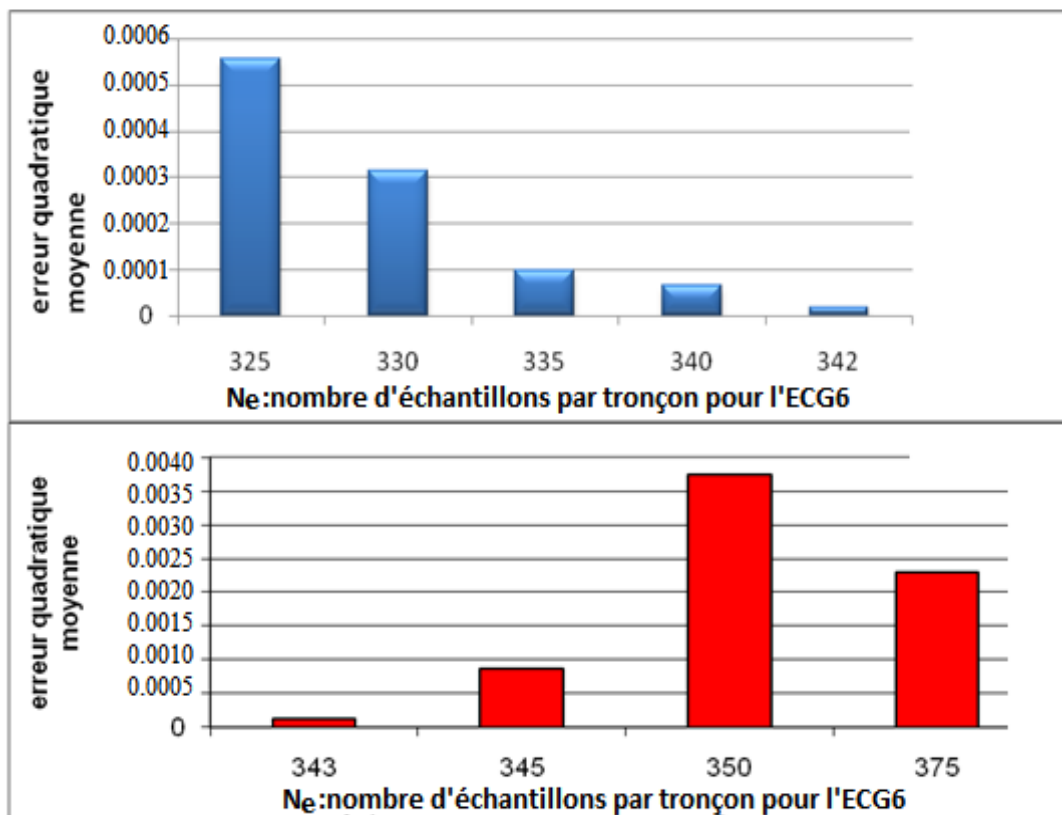


Figure 2.22 : Résultats de la variation de N_e , nombre de points par battement pour l'ECG 6.

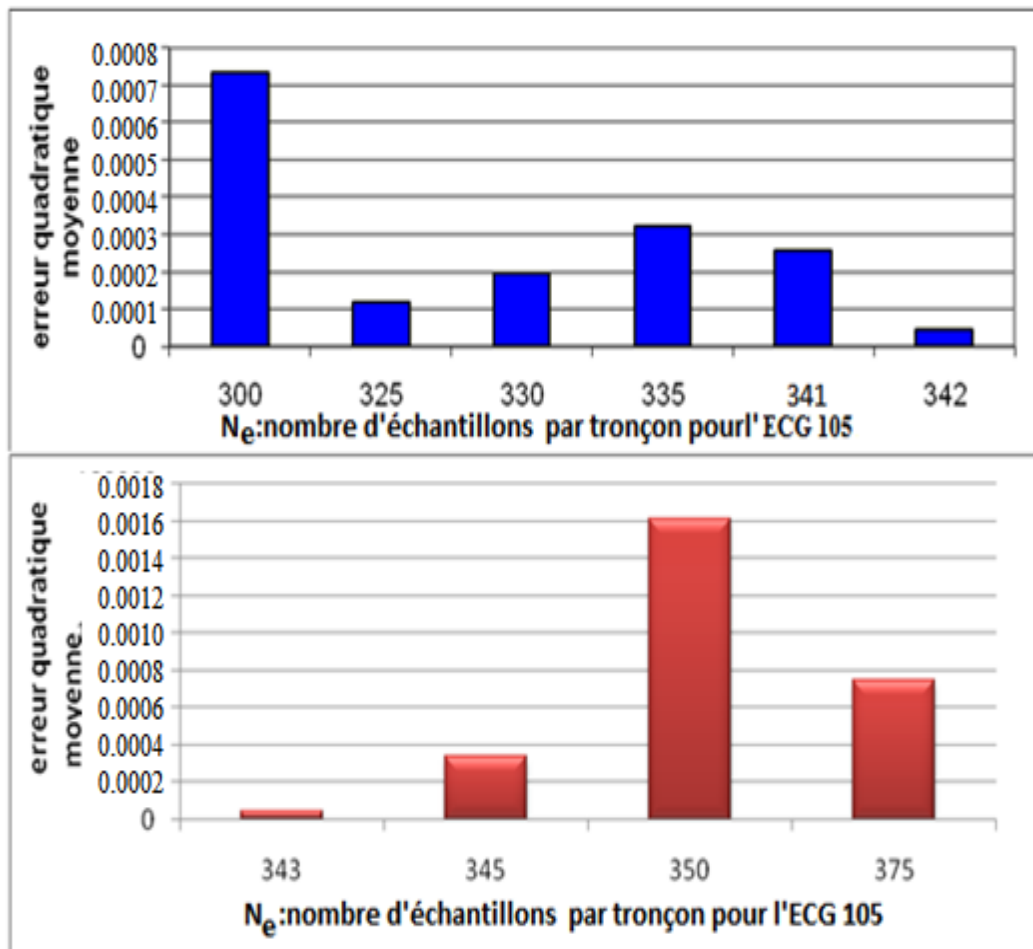


Figure 2.23 : Résultats de la variation de N_e , par tronçon pour l'ECG 105.

D'après les résultats illustrés ci-dessus (figures 2.22 et 2.23), nous avons constaté que l'erreur augmente rapidement lorsque $N_e \geq 343$; ce qui explique la déformation du signal modèle de la figure 2.24(c).

Pour l'ECG 6 : le découpage effectué avec $N_e=342$ présente une erreur quadratique moyenne de $j=0,00001721$; cette erreur est minimale (voir figure 2.22).

Pour l'ECG 105 : le découpage effectué avec $N_e=342$ donne une erreur quadratique moyenne de $j=0,000051914$; cette erreur est minimale (voir figure 2.23).

Les trois graphes des figure 2.24(a), 2.24(b), 2.24(c) représentent la comparaison entre les tronçons de l'ECG 6 réel qui est en rouge , ceux du modèle en bleu résultant d'un découpage avec un nombre de points $N_e=342$ et ceux du modèle en orange résultant d'un découpage avec $N_e=350$.

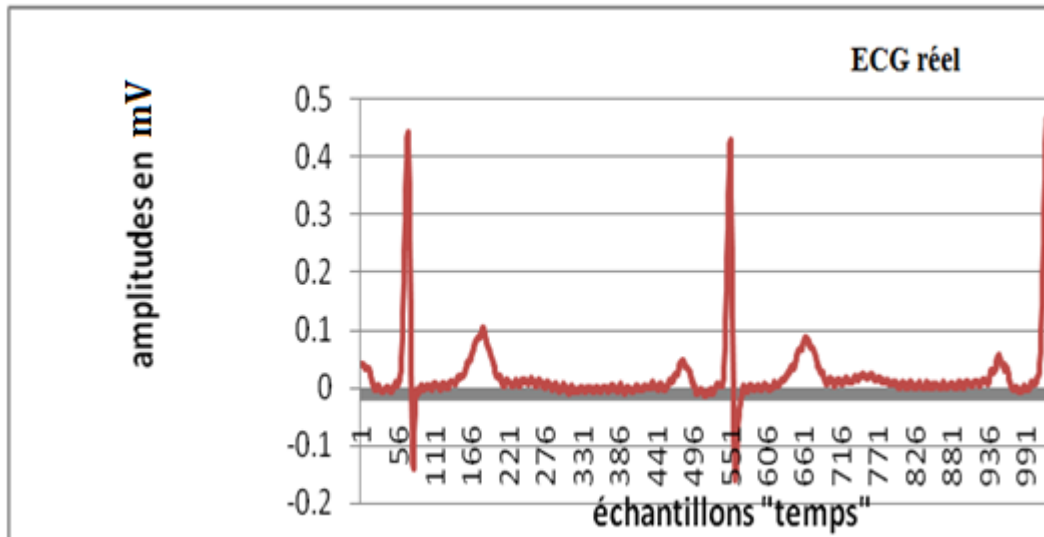


Figure 2.24(a) : Tracé de de l'ECG 6 réel.

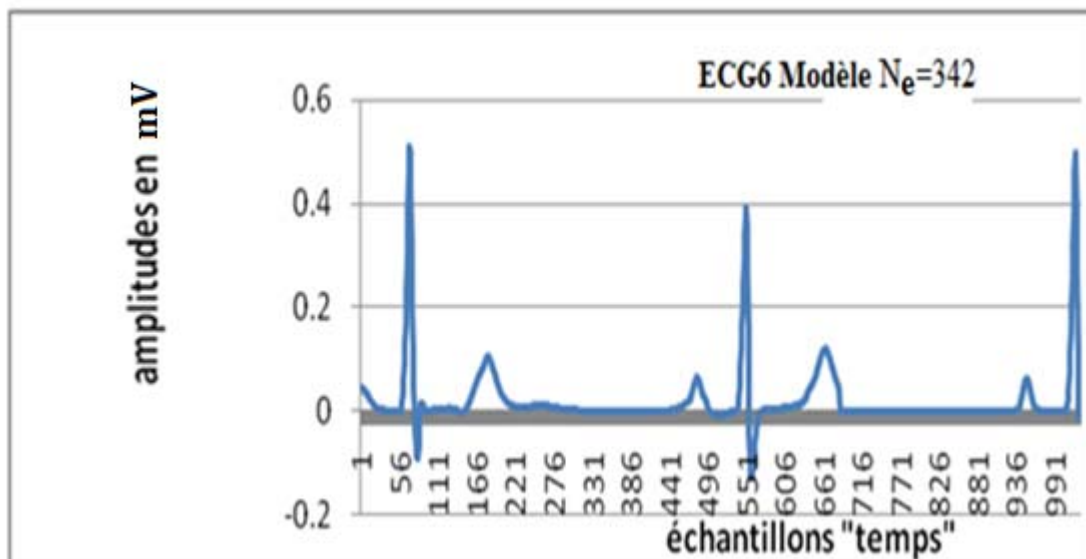


Figure 2.24(b) : Reconstitution de l'ECG 6 avec $N_e=342$.

Le modèle de la figure 2.24(c) en orange ne ressemble pas du tout au signal original de la figure 2.24(a) (en rouge) et ses valeurs sont aberrantes. Donc pour $N_e \geq 342$ le signal est complètement déformé ; cela correspond tout à fait à l'erreur qui est de l'ordre de 10^{-4} .

Par contre le modèle de la figure 2.24(b) en bleu montre que le choix de $N_e=342$ est justifié.

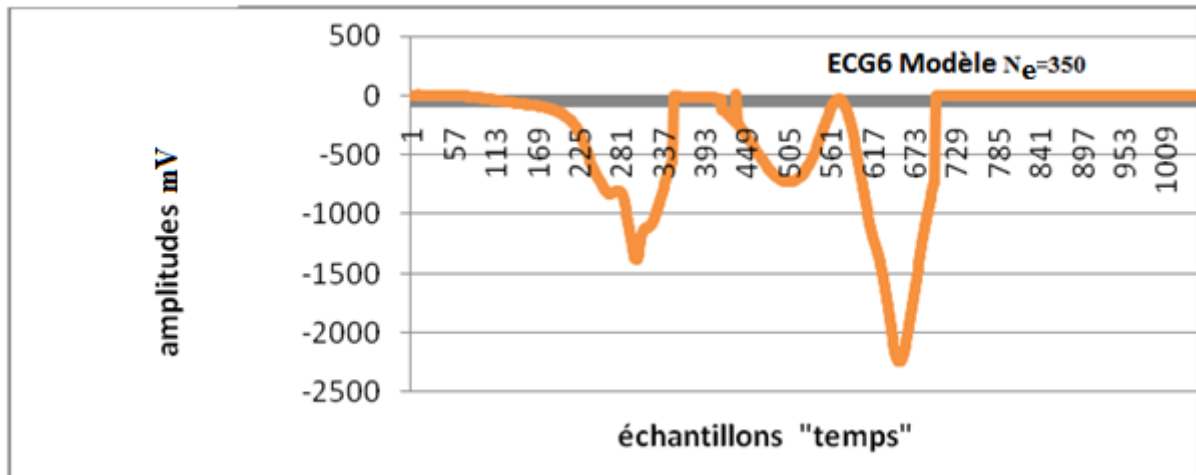


Figure 2.24(c) : Reconstitution de l'ECG 6 avec $N_e=350$.

Conclusion

Dans ce chapitre nous avons décrit en détail la modélisation du signal ECG par les réseaux de neurones RBF ainsi que toutes les étapes de l'application de l'algorithme GOFR : sélection, optimisation et orthogonalisation.

Notre réseau de neurones RBF est constitué de $N=10$ neurones dans la couche cachée. L'algorithme GOFR a pour but d'initialiser les paramètres de la fonction d'activation des N neurones. Autrement dit, le GOFR a dix gaussiennes à sélectionner afin d'initialiser les paramètres du réseau. Notre algorithme GOFR ne sélectionne pas 10 gaussiennes systématiquement, il sélectionne uniquement celles dont il a besoin pour approcher au mieux les ondes caractéristiques du tronçon à modéliser. Par contre, s'il a besoin de plus de N gaussiennes pour approcher le signal avec une erreur minimale, l'algorithme GOFR est forcé d'arrêter à la N ème sélection et ne peut la dépasser en aucun cas. C'est pour cela que le choix de N_e lors du découpage est très important ainsi que le nombre de neurones du réseau N .

Le choix judicieux de $N_e=342$ et $N=10$ est, encore une fois, justifié car l'initialisation avec l'algorithme GOFR nous a rapproché de la solution optimale. Nous l'avons constaté à travers les valeurs des erreurs commises sur chaque battement qui sont minimales de l'ordre de 10^{-5} .

Nous allons présenter au chapitre trois suivant un outil permettant la compression d'un signal ECG en garantissant la possibilité de le reconstruire. Ce dernier reposera sur un algorithme de décomposition et de reconstitution de signaux ECG à base d'ondelettes.

Introduction

Les signaux biomédicaux tels que les électrocardiogrammes (ECG) sont aujourd'hui, des données accessibles grâce à des systèmes électroniques dont les performances dépendent des techniques élaborées de traitement de signal en temps réel.

Pour permettre aux praticiens d'effectuer efficacement leurs diagnostics, des enregistrements de longue durée (24 Heures grâce au Holter.) des ECG sont nécessaires. Le problème majeur réside dans la quantité d'informations à analyser. Seuls des outils performants d'aide au diagnostic pourront apporter une solution.

L'intérêt de la compression avec reconstitution appliquée à l'ECG est primordial car techniquement on pourra améliorer la transmission et le stockage de ce signal d'une part et d'autre part cette amélioration sera bénéfique sur le plan de la pratique médicale, en raison de l'importance de l'interprétation de l'ECG pour établir un diagnostic sûr et précis.

Lors de nos travaux, nous avons élaboré une méthode permettant la compression avec reconstitution du signal électrocardiographique (ECG), afin de répondre à une forte demande du milieu médical (médecine ambulatoire et télémedecine). En compressant les signaux ECG, la possibilité de stocker un maximum de données issues des monitorings permanents dans des espaces réduits et de faciliter leurs transmissions à travers différents réseaux peut être établie. Ceci suggère une optimisation de cette compression. C'est-à-dire, qu'il faudra concilier minimisation des données, minimisation des pertes d'information et maximisation de la qualité de transmission. La mise au point et implémentation de méthodes de compression avec reconstitution des signaux ECG, est basée sur une méthode temps-fréquence : la théorie des ondelettes. Par conséquent, ce chapitre sera dédié à la présentation des outils mathématiques d'analyse fréquentielle, temps-fréquence et à la mise en évidence de leurs avantages et inconvénients, cela afin de mieux nous orienter vers l'outil permettant la meilleure compression du signal ECG.

Partie A : La transformée en ondelette (wavelet transform : WT)

Le signal ECG est un signal biologique non stationnaire ou variant. Afin de pouvoir l'interpréter au mieux, il est nécessaire d'avoir des informations de fréquence et de temps au même moment. L'utilisation d'une transformée adaptée est donc primordiale.

3.1 Passage de la transformée de Fourier à la transformée en ondelette

La transformée de Fourier (TF) ne donne pas d'informations fréquentielles et temporelles simultanément. Donc, elle ne s'applique qu'aux signaux stationnaires car leurs composantes en fréquence existent en permanence.

Pour pallier à ce problème, on applique la transformée de Fourier fenêtrée (STFT : Short Term Fourier Transform) en considérant le signal ECG comme une succession de signaux stationnaires sur des intervalles bien définis. Lorsque ces intervalles sont très petits, nous regardons le signal à travers une étroite fenêtre. Vu de cette fenêtre, le signal est indéniablement stationnaire. Ainsi, la STFT nous fournit une représentation temps-fréquence du signal. Mais elle ne satisfait pas pleinement notre critère. Tout ce que nous pouvons connaître, ce sont les intervalles de temps pendant lesquels une certaine bande de fréquences existe. Donc la STFT pose un problème de résolution dépendant de la largeur de la fenêtre.

Comme la fonction de fenêtrage est étroite (elle est à support compact) on aura une bonne résolution temporelle et une médiocre résolution fréquentielle, alors qu'on avait l'inverse pour la TF [65].

Dans cette première partie nous avons souligné brièvement les limites de la TF et de la STFT, d'où la nécessité d'introduire une nouvelle technique d'analyse pour une meilleure résolution temporelle aux basses fréquences. Cette technique plus appropriée à la forme du signal ECG est la transformée en ondelette que nous décrirons en détails. Nous expliquerons en particulier les transformées en ondelettes continue et discrète, et aborderons leurs notions mathématiques.

3.2 La théorie des ondelettes

Elle a pour but de fournir un ensemble de méthodes et d'algorithmes adaptés à des signaux non-stationnaires. A la différence de l'analyse de Fourier, qui repose sur une série de fonctions sinusoïdales, l'analyse par ondelettes utilise le temps et la fréquence pour représenter des signaux non stationnaires avec une plus grande efficacité ; elle constitue un outil très puissant permettant la représentation d'un signal par une somme finie de composantes à des résolutions différentes, de telle façon que chaque composante puisse être traitée avec souplesse, en fonction des objectifs de l'application. De ce fait, la représentation par ondelettes est beaucoup plus compacte et facile à mettre en œuvre.

3.2.1 Historique

Les premières ondelettes ont été introduites par Haar en 1909. Depuis, l'intérêt que suscite l'ondelette est de plus en plus grand. En effet, depuis les années trente de nombreux travaux ont été effectués, révélant de nouvelles ondelettes présentant diverses propriétés et s'appliquant dans des domaines tous aussi variés (physique, traitement du signal, analyse de l'image, compression des données, statistique et informatique). Nous citerons les ondelettes de Grossman-Morlet, de Gabor-Malvar, de Lusin, de Meyer ou encore celles de Daubechies, à savoir : Daubechies, symlets, et coiflet [66]. Avant d'appliquer la WT aux signaux ECG nous allons introduire son concept mathématique

3.2.2 Aspect mathématique de la transformée en ondelette

La transformée en ondelette, comme la transformée de Fourier, est une transformation temps-fréquence. Elle remplace la sinusoïde de la TF par une famille de translations et dilatations d'une même fonction. Cette fonction, nommée ondelette mère Ψ , est une fonction de fenêtrage, oscillante, de dimension finie [67].

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right). \quad (3.1)$$

La décomposition d'une fonction en ondelette consiste à l'écrire sous la forme d'une somme pondérée de fonctions obtenues à partir d'opérations simples effectuées sur l'ondelette mère. Ces opérations sont des translations τ (temporelles) et des dilatations s (échelles); selon qu'elles sont choisies de manière continue ou discrète, on parlera d'une transformée en ondelette continue ou discrète [68].

3.2.3 La transformée en ondelette continue CWT (Continu wavelet transform)

La transformée en ondelette est dite continue lorsque les paramètres de dilatation et de translation varient continûment. Bien entendu, dans la pratique numérique une variation continue n'existe pas, donc ces pas seront discrétisés finement.

Cette transformée est définie comme étant le produit scalaire du signal $x(t)$ et l'ondelette mère $\Psi(t)$:

$$CWT_x^\Psi(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \Psi^*\left(\frac{t-\tau}{s}\right) dt \quad (3.2)$$

s = échelle (coefficient de dilatation de l'ondelette) = $1/f$,

τ = coefficient de translation (déplacement de la fenêtre le long du signal),

Ψ^* : le conjugué de l'ondelette mère.

L'algorithme CWT réalise les étapes suivantes:

1. Choisir le type de l'ondelette mère,
2. Positionner la fonction d'ondelette d'échelle $s=1$ (les hautes fréquences) à l'origine du signal ($t=0$),
3. multiplier par le signal puis intégrer sur tous les temps,
4. multiplier le résultat par $1/\sqrt{s}$ (pour normaliser l'énergie),
5. décaler l'ondelette vers la droite par ($t=\tau$),
6. Revenir en 3, et répéter la procédure jusqu'à ce que la fenêtre atteigne l'extrémité du signal,
7. Incrémenter s puis répéter toutes les étapes citées précédemment.

La figure 3.1 illustre cette méthode de calcul.

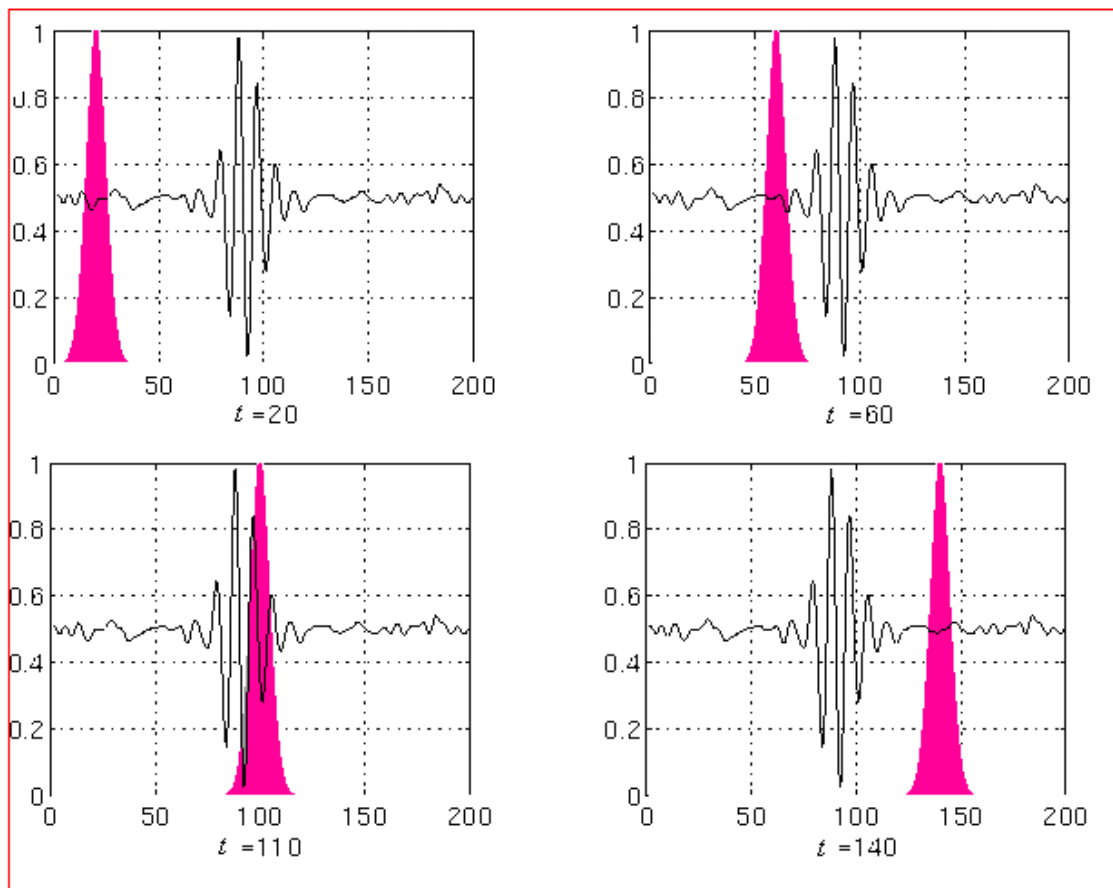


Figure 3.1: Transformée en ondelette continue à l'échelle $s=5$.

Si une composante spectrale du signal correspond à la valeur courante de s , le produit de l'ondelette par le signal, à l'emplacement où existe la composante spectrale, donne une valeur

élevée (coefficient important car le produit entre eux est positif). Si la composante spectrale correspondant à la valeur courante de s n'est pas présente dans le signal, la valeur de ce produit sera nulle ou très faible [65,69].

La figure 3.2 est une représentation temps-échelle résultant de l'application de la CWT à un signal ECG ; à chaque instant t et échelle s (scale) on a la valeur du coefficient d'ondelette, issue du produit du signal ECG analysé et de l'ondelette analysante.

Notons que l'incrémentement de τ et s se fait par petites valeurs pour assurer la continuité, c'est-à-dire qu'on effectue un échantillonnage très fin du plan temps-fréquence.

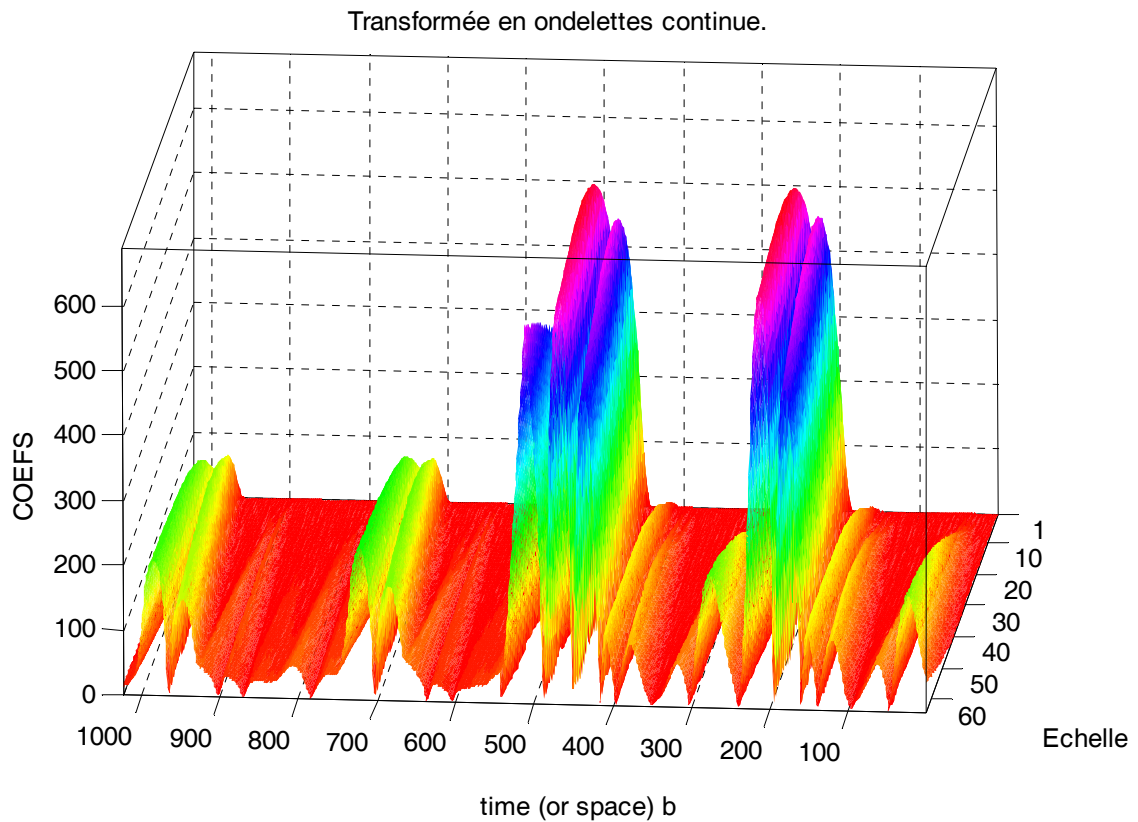


Figure 3.2 : CWT appliquée à un signal ECG présentant une ectopie ventriculaire.

Bien que la transformée en ondelettes continues discrétisées permette le calcul de la CWT par des ordinateurs, elle n'est pas une vraie transformée discrète, et les informations qu'elle fournit sont fortement redondantes. Cette redondance, d'autre part, exige une quantité significative de temps et de ressources de calcul.

3.2.4 La transformée en ondelette discrète (Discret Wavelet Transform : DWT)

Contrairement à la CWT, la DWT est simple à implémenter et assure suffisamment d'informations, tant pour l'analyse que pour la reconstitution du signal et cela en un temps notablement réduit.

3.2.4.1 Expression de la famille d'ondelettes discrètes

Dans la DWT la discrétisation du plan temps-fréquence correspond à la construction d'une grille non uniforme définie par les paramètres représentant les coefficients d'échelle s et de dilatation τ calculés de la manière suivante :

$$\begin{cases} s = s_0^i \\ \tau = k \cdot \tau_0 s_0^i \end{cases} \quad \text{avec } s_0 > 1 \text{ et } \tau_0 > 0 \quad i, k \in \mathbb{N} \quad (3.3)$$

Les ondelettes sont alors définies par :

$$\psi_{m,n}(t) = \frac{1}{\sqrt{s_0^m}} \psi\left(\frac{1}{s_0^m} t - n\tau_0\right). \quad (3.4)$$

L'idée essentielle derrière ce schéma de discrétisation est la suivante :

Une fois choisi, un grossissement (i.e. s_0) étudie le processus à une position donnée, ensuite nous nous déplaçons à une autre position. Si le grossissement est important (analyse des détails à petite échelle), les déplacements se font par petits incréments. Au contraire, si le grossissement est faible (analyse à plus grande échelle), les déplacements peuvent être plus rapides. Ceci est automatiquement accompli si l'incrément de déplacement est inversement proportionnel au grossissement. La transformée en ondelette discrète s'écrit alors :

$$DWT_x^\psi = \frac{1}{\sqrt{s_0^i}} \int x(t) \psi\left(\frac{1}{s_0^i} t - k\tau_0\right) \quad (3.5)$$

3.2.4.2 Implémentation de la transformée en ondelette discrète

Dans la transformée en ondelette discrète, les coefficients τ_0 et s_0 prennent les valeurs 1 et 2 respectivement, qui correspond à un échantillonnage dyadique du plan temps-fréquence.

Une des raisons du succès de la DWT est son implantation matérielle efficace. La méthode la plus répandue pour l'implémenter est l'utilisation de bancs de filtre sous forme

d'une structure pyramidale, technique qui provient des travaux de Mallat. Deux filtres FIR (Finite Impulse Response) h et g [Annexe A1] décomposent le signal en une approximation grossière et des informations de détail: ils possèdent alors des caractéristiques respectivement passe-bas et passe-haut. Pour garder le même nombre d'échantillons en sortie et en entrée, les produits de convolution issus des filtres sont sous-échantillonnés par un facteur deux. Seule la sortie du filtre passe-bas, c'est-à-dire l'approximation, est de nouveau traitée par les deux filtres. Cette structure pyramidale est illustrée sur la figure 3.3.

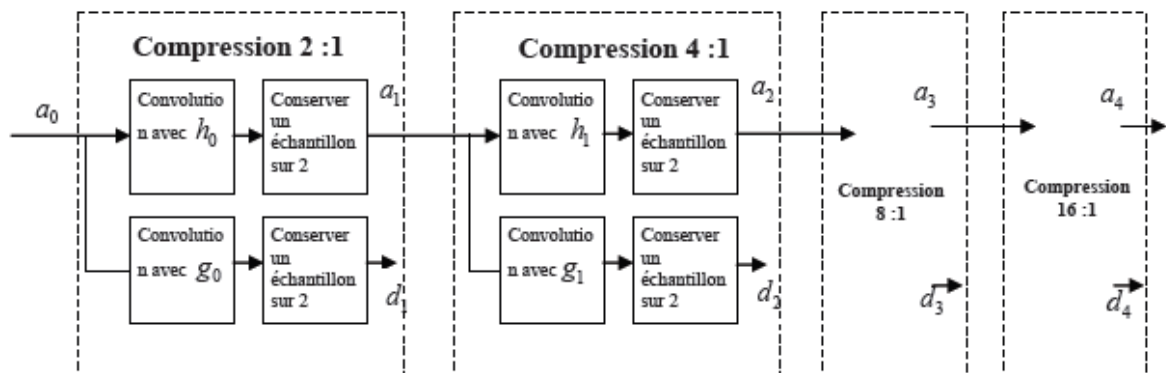


Figure 3.3 : Principe de la transformée en ondelette discrète calculée à l'aide d'un banc de filtre.

Le nombre total de coefficients (d'approximation et de détail) ainsi obtenu est égal au nombre d'échantillons contenus dans le signal. Partant de ces coefficients le signal original peut être reconstruit et une localisation temporelle des composantes fréquentielles du signal est possible [70].

La transformée inverse peut être obtenue en sur-échantillonnant le signal par insertion d'un zéro tous les deux coefficients et en utilisant une autre paire de filtres FIR, figure 3.4. Les quatre filtres (deux pour la décomposition et deux pour la reconstruction) forment un banc de filtres conjugués en quadrature [71].

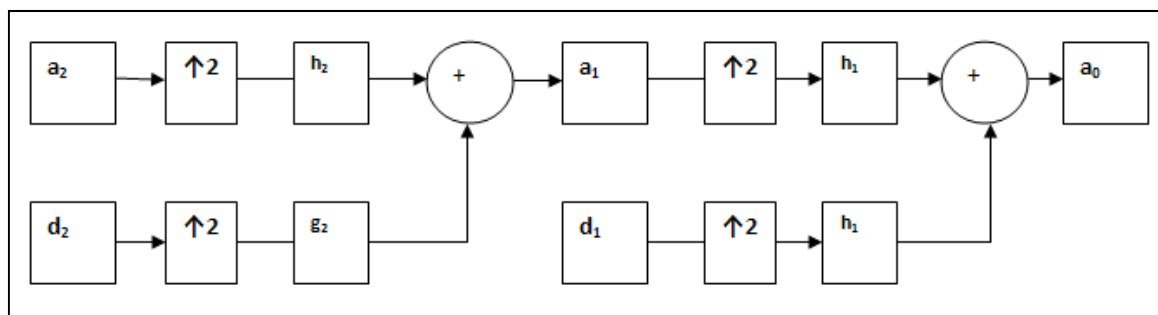


Figure 3.4 : Reconstitution à base d'ondelettes avec filtre FIR.

La DWT représente ainsi le signal avec le minimum de coefficients nécessaires, soit autant de coefficients que de points d'échantillonnage, sans pertes d'informations : pour les basses fréquences, moins de coefficients que pour les hautes fréquences sont obtenus, avec un rapport de puissance de 2 entre chaque étape de décomposition. Donc la DWT ,par exploitation du nombre de coefficients d'ondelette exactement nécessaire, n'est envisageable qu'à la condition que les ondelettes soient orthogonales.

3.2.5. L'orthogonalité

L'orthogonalité signifie que l'information capturée par une ondelette est totalement décorrélée de celle capturée par une autre. Cela permet de ne garder que l'information nécessaire et suffisante afin d'assurer la réversibilité. Une ondelette mère orthogonale et sa fonction d'échelle permettent d'effectuer une analyse multi-résolution orthogonale : les espaces de détail et d'approximation sont alors orthogonaux, c'est-à-dire que la projection des vecteurs de la base de l'un des espaces sur l'autre donne zéro. Chaque ondelette et sa fonction d'échelle associée sont également orthogonales, [73]. Le spectre des ondelettes orthogonales est donné en figure 3.5.

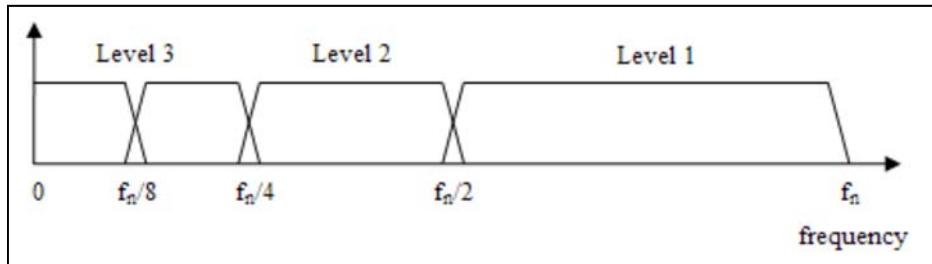


Figure 3.5: Spectre des ondelettes orthogonales.

La fonction d'ondelette Ψ représente un filtre passe haut, qui définira les détails, sa fonction correspondante aura pour expression :

$$\Psi(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} g_n[n] \Phi(2t - n) \quad (3.6)$$

Les approximations sont définies par la fonction d'échelle Φ qui représente un filtre passe bas, sa fonction correspondante est :

$$\Phi(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} h_n[n] \Phi(2t - n) \quad (3.7)$$

3.2.6. Application de la décomposition par DWT à un signal ECG

Les signaux ECG sur lesquels nous faisons nos tests proviennent de la MIT-BIH Data base, [74,75]. Ils sont récapitulés dans le tableau suivant :

N° des ECG de la MIT Database	Type d'ECG	Fréquence d'échantillonnage
16273	Normal sinus rythme (Sain)	360Hz
cu19	Tachyarythmie ventriculaire	360Hz
04015	Fibrillation atriale	360Hz
207	Ectopie supraventriculaire	360Hz
105	Ectopie ventriculaire	360Hz
cu01	Tachyarythmie ventriculaire	360Hz
X_04936	Fibrillation	360Hz
103	Normal sinus rythme (Sain)	360Hz

Nous avons choisi d'effectuer une décomposition des signaux ECG sur trois niveaux (N=3) en utilisant l'ondelette de Haar définie par :

$$\psi(x) = \begin{cases} 1, & \text{si } x \in [0, 0.5[\\ -1, & \text{si } x \in [0.5, 1[\end{cases} \quad \text{et} \quad \psi(x) = 0, \text{ si } x \notin [0, 1[$$

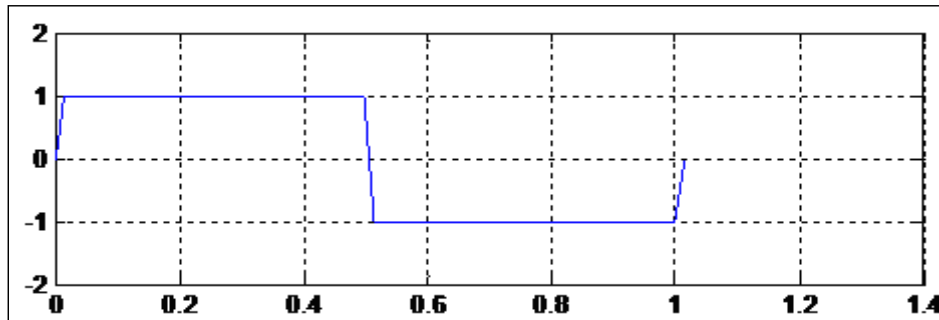


Figure 3.6: Tracé de la fonction d'ondelette de Haar.

L'ondelette de Haar de la figure 3.6 est une ondelette orthogonale permettant le calcul des deux types de transformée en ondelette, continue et discrète, [76]. L'impact de la propriété d'orthogonalité dans notre application sera expliqué ultérieurement.

Nous procédons à la décomposition d'un extrait de 1024 échantillons du signal ECG 105 de la figure 3.7.

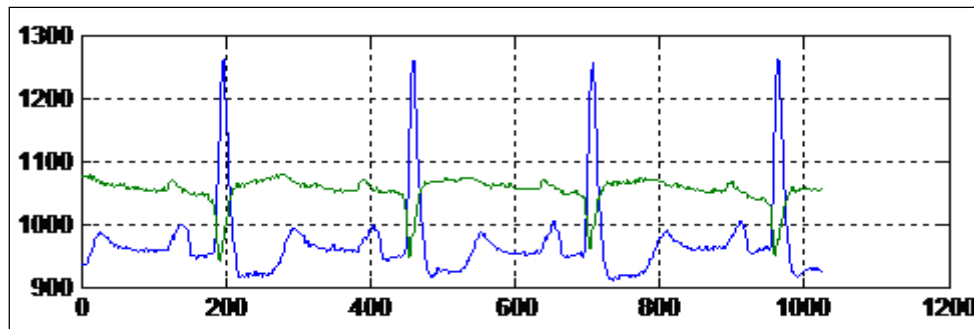


Figure 3.7 : Signal ECG 105.

Cet ECG correspond à une patiente, âgée de 73 ans, présentant une ectopie ventriculaire.

Les coefficients d'approximations et de détails, résultants de la décomposition d'une des dérivations, sont représentés en figure 3.8.

On remarque que le tracé des coefficients d'approximation au troisième niveau donne l'allure générale du signal traité, alors que le tracé des coefficients de détails est comparable à une action de « Zoom » sur le signal.

D'autre part, si l'on observe l'axe des abscisses correspondant aux coefficients d'approximations et ceux de détails des trois niveaux, on trouve que le nombre total des coefficients est égale à 1024 qui est le nombre d'échantillons du signal original.

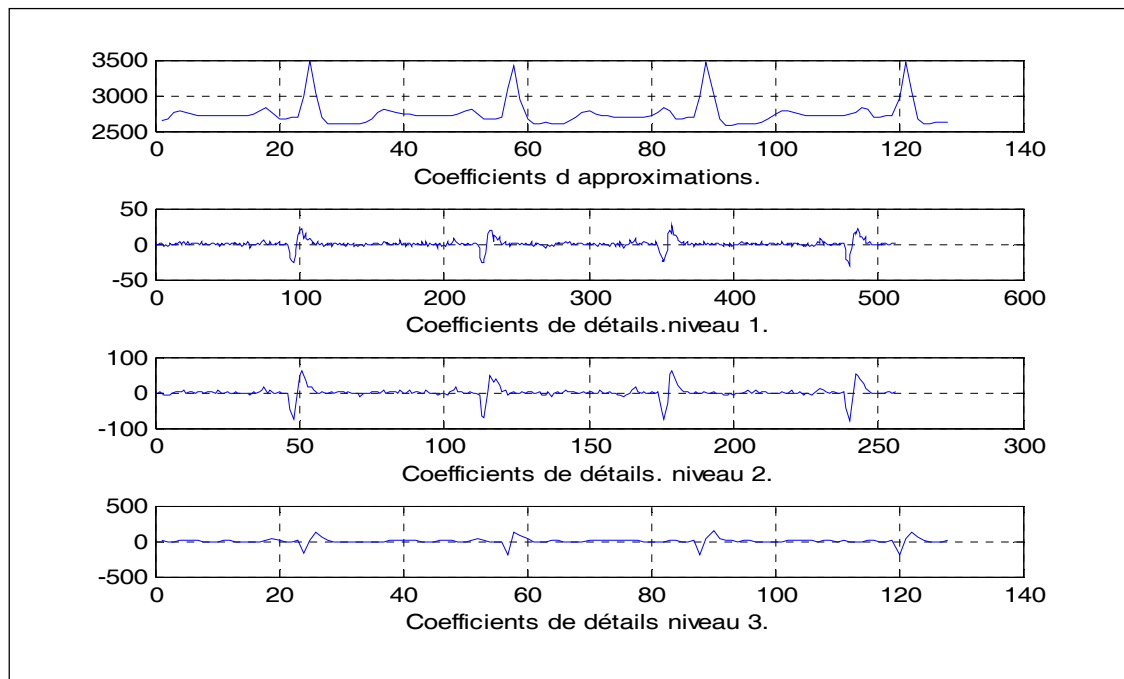


Figure 3.8 : Tracé des coefficients d'approximation et de détails.

3.2.7. Application de la CWT et de la DWT au signal ECG

Nous présentons dans ce qui suit les temps de calcul obtenus lors de la décomposition des signaux ECG, provenant de la base de donnée MIT BIH Database (105,103,207,217), par application de la CWT et DWT aux niveaux maximaux de décomposition en utilisant soixante-deux ondelettes fournies par le logiciel Matlab7.1.

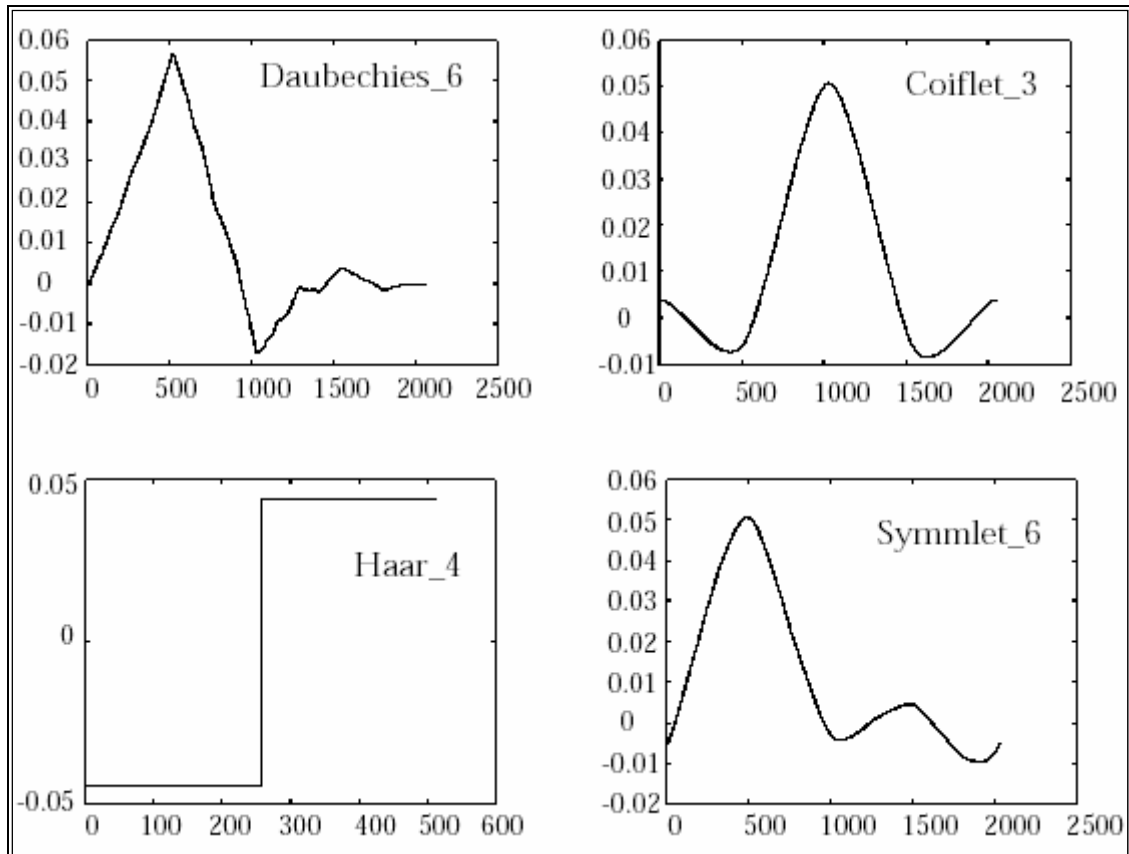


Figure 3.9 : Tracé de quelques ondelettes.

Les tracés de la figure 3.9 représentent les allures de quatre ondelettes mères, présentant différentes caractéristiques mais ayant toutefois un point commun : leur support compact.

L'ondelette de Haar est la plus simple d'entre elles, son expression a été donnée au paragraphe 6. Elle est caractérisée par son antisymétrie.

Les ondelettes de Daubechies, notées dbN , n'ont pas d'expressions explicites à l'exception de l'ondelette $db1$, qui est l'ondelette de Haar. La majorité des dbN sont asymétriques.

Les ondelettes Symlet sont des ondelettes de Daubechies, qui sont modifiées afin d'augmenter leur symétrie.

Les ondelettes Coiflet, notées *coifN*, ont été construites par Daubechies à la demande de Coifman. Elles sont beaucoup plus symétriques que les *dbN* [76].

La liste des ondelettes utilisées est donnée dans le tableau3.1 ci-dessous :

Tableau3.1 : Tableau* récapitulatif des noms des ondelettes utilisées.

N°	Ondelettes	N°	Ondelettes	N°	Ondelettes	N°	Ondelettes
1	gauss8	17	coiflet3	33	biorthogona2,8	49	reverse biorth1,3
2	Mexican hat	18	coiflet4	34	biorthogona3,1	50	reverse biorth1,5
3	morl	19	coiflet5	35	biorthogona3,3	51	reverse biorth2,2
4	daubechies1	20	symlet2	36	biorthogona3,5	52	reverse biorth2,4
5	haar	21	symlet3	37	biorthogona3,7	53	reverse biorth2,6
6	daubechies2	22	symlet4	38	biorthogona3,9	54	reverse biorth2,8
7	daubechies3	23	symlet5	39	biorthogona4,4	55	reverse biorth3,1
8	daubechies4	24	symlet6	40	biorthogona5,5	56	reverse biorth3,3
9	daubechies5	25	symlet7	41	biorthogona6,8	57	reverse biorth3,5
10	daubechies6	26	symlet8	42	complexgau8	58	reverse biorth3,7
11	daubechies7	27	biorthogonal1,1	43	complexmor1.5-1	59	reverse biorth3,9
12	daubechies8	28	biorthogonal1,3	44	shanon0.5-1	60	reverse biorth4,4
13	daubechies9	29	biorthogona1,5	45	fb spline2-0.5-1	61	reverse biorth5,5
14	daubechies10	30	biorthogona2,2	46	meyer	62	reverse biorth6,8
15	coiflet1	31	biorthogona2,4	47	discret meyer		
16	coiflet2	32	biorthogona2,6	48	reverse biorth1,1		

* Pour connaître les propriétés de ces ondelettes se référer à l'annexe A2.

Les résultats obtenus sont donnés sous forme de graphiques de la figure 3.10 à 3.12.

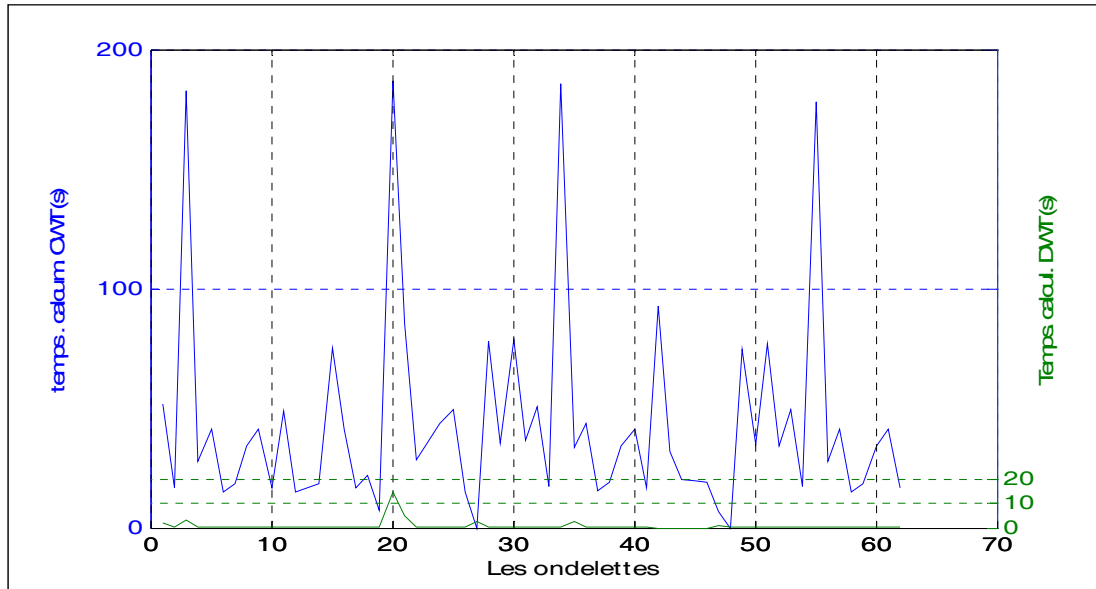


Figure 3.10 : Tracé des temps de décomposition de l'ECG 103 par CWT et DWT.

On remarque que les temps de calcul correspondant aux transformées en ondelettes continues sont très élevés par rapport à ceux correspondant aux transformées en ondelettes discrètes (jusqu'à 100 fois plus élevés) et cela pour la quasi-totalité des ondelettes.

En effet, pour les ondelettes 27 et 47 (bior1.1 et rbio1.3) nous n'avons pas pu appliquer la CWT à cause de l'insuffisance de l'espace mémoire, car elles engendraient un nombre important de coefficients, contrairement à la DWT. Cela traduit la redondance de la CWT.

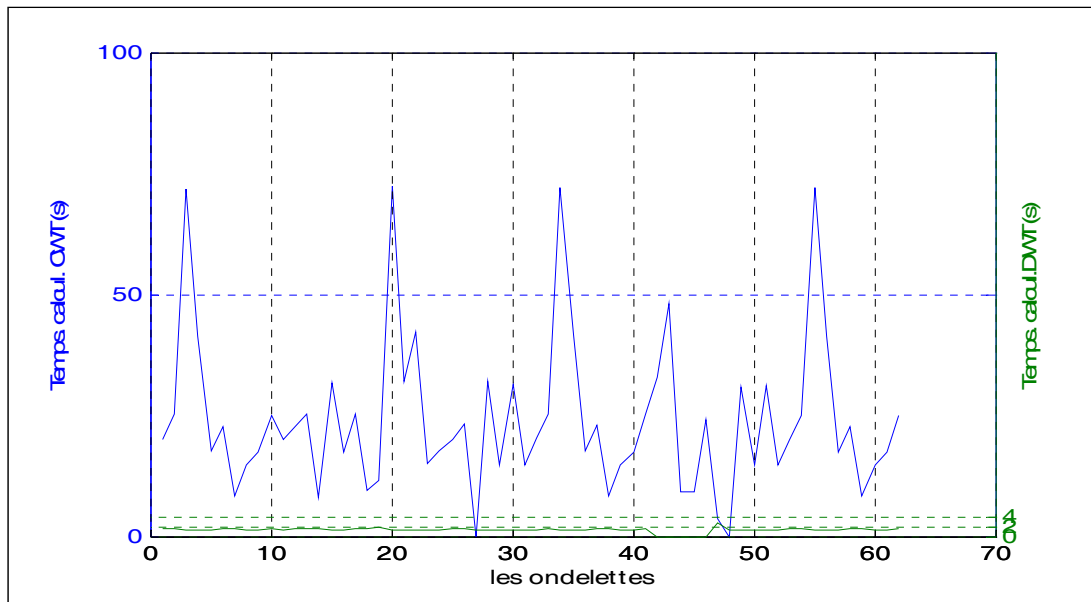


Figure 3.11: Tracé des temps de décomposition de l'ECG 105 par CWT et DWT.

D'autre part, nous avons constaté que pour les ondelettes de N° 1, 2 et 3 et celles allant du N° 42 à 46, la DWT est inapplicable, car elles sont réservées à la CWT.

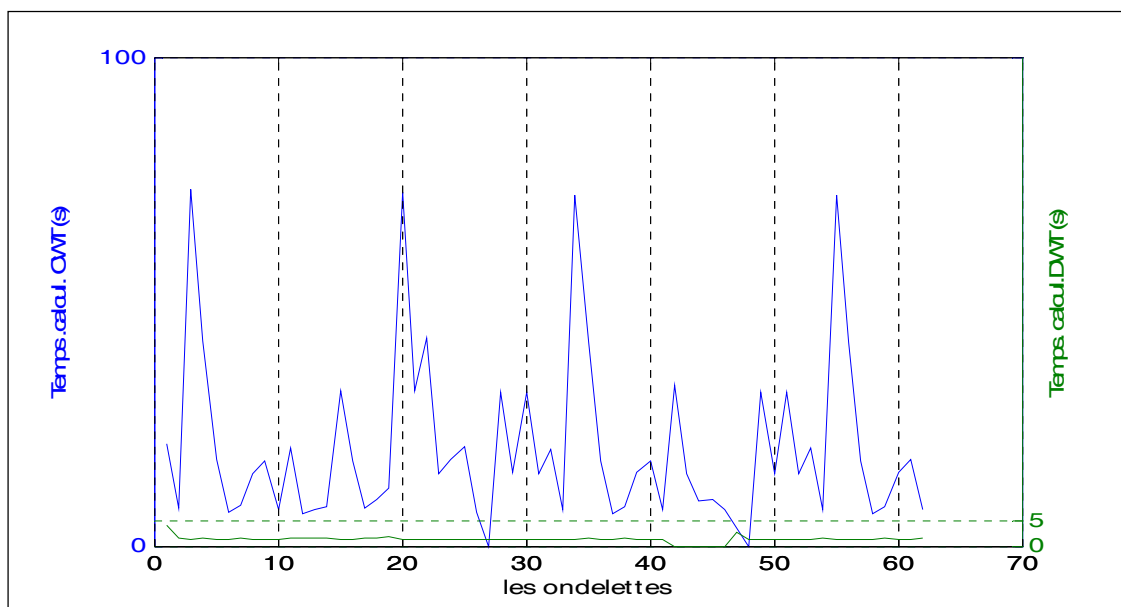


Figure 3.12 : Tracé des temps de décomposition de l'ECG 207 par CWT et DWT.

Dans cette partie A, les résultats obtenus mettent en évidence la différence majeure existant entre les deux types de transformées. En effet, la CWT engendre des temps de calcul

colossaux en comparaison avec la DWT du fait du grand nombre de coefficients qu'elle doit calculer. Indéniablement la CWT est redondante, ce qui la rend d'une part coûteuse en termes de temps et de capacité de stockage et incommode d'autre part puisqu'elle ne permet pas la reconstitution du signal.

Notre but qui est la compression du signal ECG, exige d'avoir un maximum d'informations avec un minimum de taille mémoire en gardant la possibilité de reconstruction. Cela nous conduit à baser la suite de notre étude sur la DWT afin de trouver l'ondelette permettant la meilleure reconstitution en un temps notablement réduit.

Partie B : Choix de l'ondelette

Ce choix dépend du type d'analyse à réaliser et du signal pris en compte, Dans notre cas l'ondelette est utilisée comme un outil permettant la compression d'un signal ECG en garantissant la possibilité de le reconstruire. Dans ce but nous nous sommes intéressées aux ondelettes discrètes.

Mais avant toute chose nous devons introduire la notion de reconstitution après décomposition dans le cadre de la transformée en ondelette.

3.3 Condition de reconstitution du signal

Le signal x peut être reconstruit à partir de la formule qui suit:

$$x(t) = \frac{1}{c_\psi} \iint_R \Psi_x^\psi(\tau, s) \psi \frac{1}{\sqrt{s}} \left(\frac{t-\tau}{s} \right) d\tau ds \quad (3.8)$$

à condition que l'intégrale suivante soit convergente :

$$c_\psi = \int_0^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega \quad (3.9)$$

($\hat{\psi}$ est la transformée de Fourier de l'ondelette mère ψ).

Cette dernière condition est aussi appelée critère d'admissibilité ; elle est intéressante dans la mesure où elle donne des informations sur les propriétés que doit vérifier une ondelette mère si l'on souhaite reconstruire le signal d'origine.

En particulier, nous devons avoir: $\hat{\psi}[\omega]_{\omega=0} = 0$. Cela induit deux conséquences importantes :

1. les ondelettes doivent posséder un spectre de type passe-bande,
2. en réécrivant l'équation précédente sous la forme : $\int_{-\infty}^{+\infty} \psi(t) dt = 0$ on s'aperçoit que $\psi(t)$ doit être à moyenne nulle [68].

3.4 Décomposition et reconstitution pratique du signal ECG

Dans cette application, nous avons conçu et programmé un algorithme de décomposition et de reconstitution de signaux ECG à base d'ondelettes.

Les cinquante-deux (52) ondelettes utilisées, telles que les ondelettes de Daubechies, symmlet ou bien les ondelettes biorthogonales, qui sont issues du logiciel Matlab7.1, satisfont à la condition de reconstitution

L'estimation des performances de reconstitution après chaque niveau de décomposition, pour les différentes ondelettes, se fait sur la base du calcul de l'erreur quadratique moyenne.

Cette erreur est définie par :

$$EQM = \frac{1}{n} \sum_{i=1}^n (ecg_i - ecgr_i)^2 \quad (3.10)$$

où ecg représente le signal original et $ecgr$ le signal reconstitué.

Afin de déterminer l'éventuelle relation entre niveau de décomposition et performance de reconstitution, nous avons représenté les erreurs obtenues en fonction des niveaux de décomposition pour chacune des ondelettes analysantes pour différents signaux ECG. Les deux courbes de la figure 3.13 (Erreur de reconstitution en fonction du niveau de décomposition) associées aux deux dérivation, présentes dans l'enregistrement fourni par la base de données normal sinus rythme de MIT-BIH, montrent une augmentation quasi proportionnelle. En effet, nous remarquons que plus on décompose le signal plus l'erreur de reconstitution augmente mais avec une certaine variabilité ; si l'on observe les deux dérivation on constate une augmentation nettement plus accrue pour la deuxième dérivation.

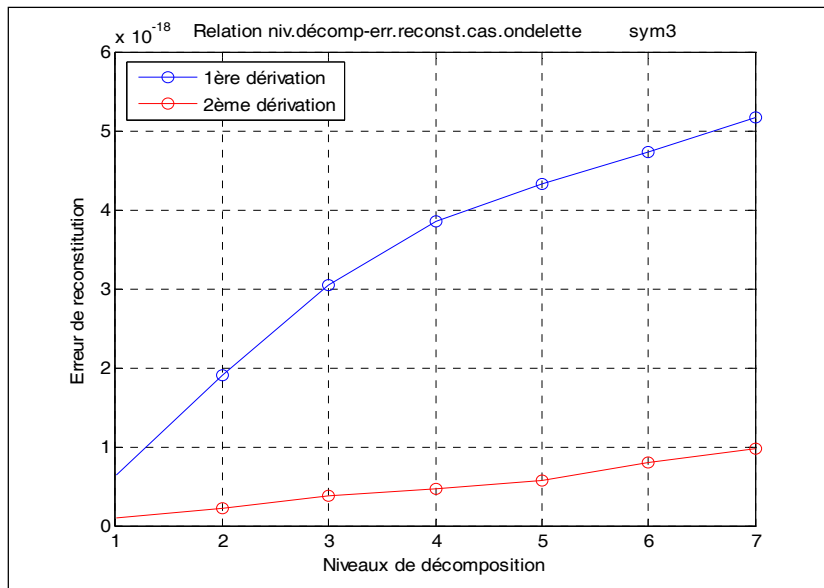


Figure 3.13 : Cas de l'ECG 16273 d'un sujet sain.

Comme dans le cas précédent, dans la figure 3.14, on constate que l'erreur de reconstitution augmente avec l'augmentation du niveau de décomposition du signal. Or, dans le cas présent, les deux dérivation suivent la même progression. Ce point de divergence est dû à la spécificité du signal traité.

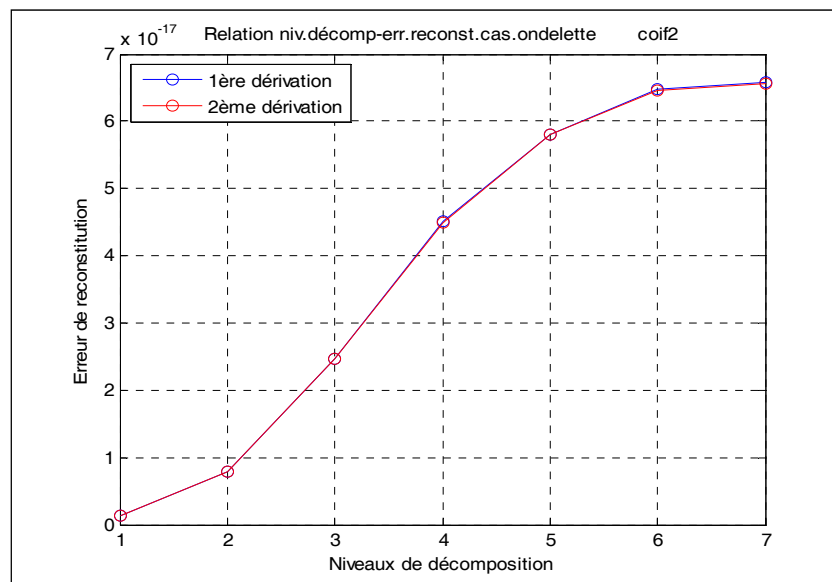


Figure 3.14 : Cas de l'ECG cu19 d'un patient souffrant de tachycardie ventriculaire.

En effet, d'une pathologie à une autre les deux dérivation utilisées peuvent avoir des allures plus ou moins similaires. Ainsi l'erreur calculée sera plus ou moins identique. Pour

l'enregistrement de la figure 3.15, nous remarquons que l'erreur résultante de la deuxième dérivation est plus élevée que la première.

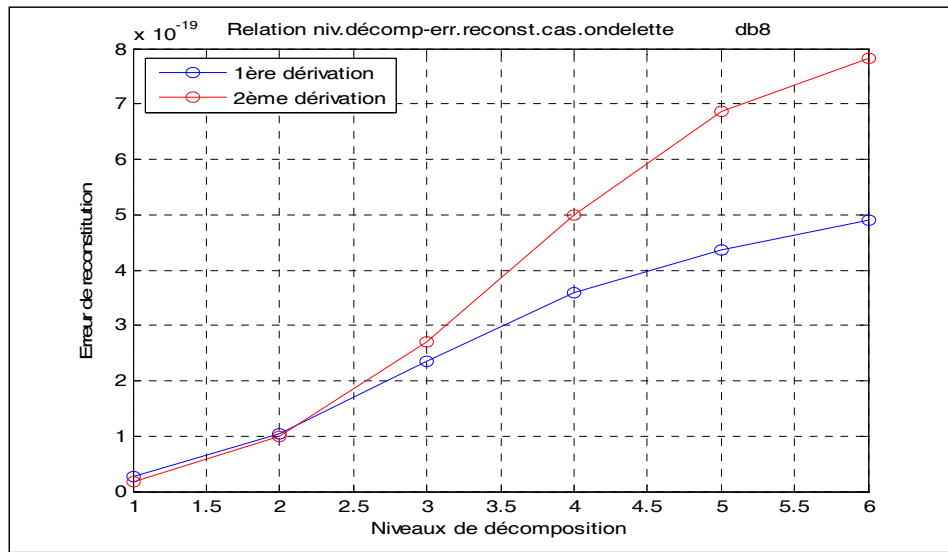


Figure 3.15 : Cas de l'ECG 04015 d'un patient souffrant de fibrillation atriale.

Dans le cas de l'ECG de la figure 3.16, moyennant l'ondelette *bior3.9*, nous constatons des erreurs identiques pour les deux dérivation. Cela est intéressant si l'on envisage d'effectuer une décomposition de ces deux signaux avec la même ondelette.

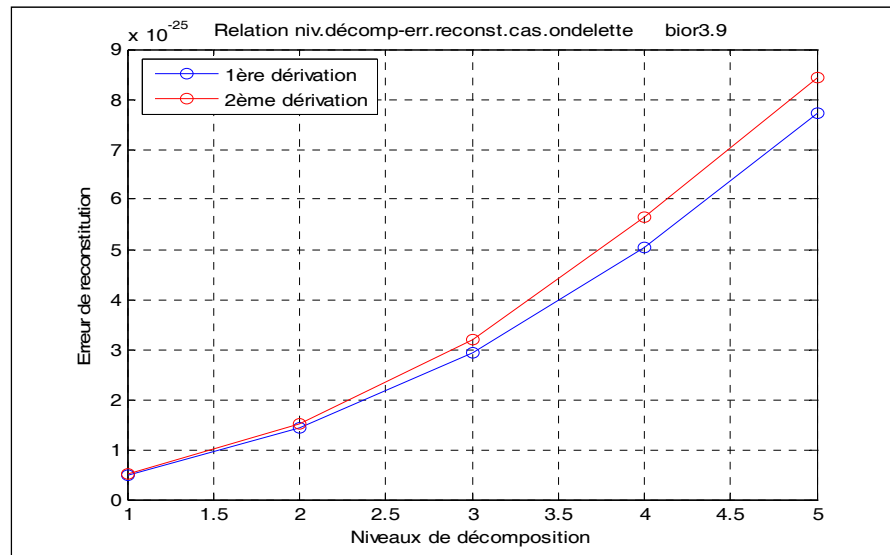


Figure 3.16 : Cas de l'ECG 207 d'un patient souffrant d'ectopie supraventriculaire.

Finalement, nous avons pu voir que l'erreur de reconstitution augmente avec les niveaux de décomposition pour les deux dérivation contenues dans les enregistrements, quelle que soit l'ondelette analysante. Si nous nous arrêtons là, le fait de prendre un niveau de

décomposition bas serait judicieux, mais il ne faut pas perdre de vue l'objectif, qui est la compression. Donc, avant de fixer un niveau il faut tenir compte du taux de compression.

Nous avons aussi voulu, à partir des algorithmes que nous avons réalisés et programmés, déterminer les ondelettes ayant les erreurs de reconstitution minimales et ceci pour chaque niveau de décomposition.

Les tableaux des figures 3.2, 3.3 et 3.4 montrent les résultats obtenus pour chaque enregistrement tiré de la base de données d'Arrhythmia MIT-BIH dans le cas des deux dérivations.

Enregistrement 103:

Tableau3.2 Ondelettes les plus performantes pour l'ECG 103.

<i>Première dérivation</i>			<i>Deuxième dérivation</i>		
Niveau de décomposition	Erreur minimale	Ondelette analysante	Niveau de décomposition	Erreur minimale	Ondelette analysante
1	1,82E-26	bior2, 2	1	1,85E-26	bior2, 2
2	4,34E-26	bior2, 2	2	4,59E-26	bior2, 2
3	8,41E-26	bior2, 2	3	9,28E-26	bior2, 4
4	1,36E-25	bior2, 2	4	1,51E-25	bior2, 2
5	1,96E-25	bior2, 2	5	2,24E-25	bior2, 4
6	0	sym8	6	2,92E-25	bior2, 4
7	9,57E-25	Cgau8	7	1,06E-24	bior1, 1
8	4,10E-24	bior2, 8	8	4,68E-24	bior2, 8

Nous pouvons voir à partir du tableau 3.2, que les erreurs de reconstitution sont très faibles pour tous les niveaux de décomposition, l'erreur minimale est de $1,82 \cdot 10^{-26}$ pour la première dérivation et de $1,85 \cdot 10^{-26}$ pour la seconde.

En réalité, la comparaison ne peut se faire que pour les premiers niveaux, car l'ensemble des ondelettes utilisées n'atteignent pas le niveau $N_{max}=8$, le niveau maximal de décomposition, qui est calculé en fonction de l'ondelette analysante et du signal ECG analysé. Ce niveau maximal varie dans le cas de l'ECG 103 entre $N_{max}=3$ et $N_{max}=8$. Les ondelettes qui donnent les meilleurs résultats sont : la biorthogonale 2.2 et la biorthogonale 2.4.

Pour l'ECG 105 (tableau 3.2), les meilleurs résultats sont obtenus avec la biorthogonale 2.2 et la biorthogonale 2.6. Les niveaux de décomposition varient entre $N_{max}=3$ et $N_{max}=6$ et les erreurs restent de très faibles valeurs

Enregistrement 105:

Tableau 3.3: Ondelettes les plus performantes de l'ECG 105.

<i>Première dérivation</i>			<i>Deuxième dérivation</i>		
Niveau de décomposition	Erreur minimale	Ondelette analysante	Niveau de décomposition	Erreur minimale	Ondelette analysante
1	1.91 E-26	bior2, 2	1	1.78 E-26	bior2, 2
2	4.43 E-26	bior2, 2	2	4.92 E-26	bior2, 2
3	8.42 E-26	bior2, 2	3	9.98 E-26	bior2, 6
4	0	bior6, 8	4	1.60 E-25	bior2, 6
5	0	coif5	5	2.33 E-25	bior2, 6
6	0	mexh	6	3.23E-25	bior2, 6

L'exemple de l'ECG 207 va dans le même sens des résultats obtenus avec les deux autres enregistrements d'ECG.

Enregistrement 207:

Tableau 3.4 : des ondelettes les plus performantes de l'ECG 207

<i>Première dérivation</i>			<i>Deuxième dérivation</i>		
Niveau de décomposition	Erreur minimale	Ondelette analysante	Niveau de décomposition	Erreur minimale	Ondelette analysante
1	1,66E-26	bior2, 2	1	1,73E-26	bior2, 2

2	4,40E-26	bior2, 2	2	4,89E-26	bior2, 2
3	8,32E-26	bior2, 2	3	9,78E-26	bior2, 2
4	1,40E-25	bior2, 2	4	1,62E-25	bior2, 2
5	1,99E-25	bior2, 2	5	2,38E-25	bior2, 2
6	2,89E-25	bior2, 2	6	3,43E-25	bior2, 2
7	3,60E-24	bior2, 8	7	3,69E-24	bior2, 8

En conclusion ; ces résultats nous orientent vers le choix de l'ondelette que nous allons utiliser plus tard dans la compression du signal ECG. Nous pouvons déjà dire que le choix se portera sur les ondelettes biorthogonales puisque, comme le montrent les tableaux précédents, leurs erreurs de reconstitution sont les plus satisfaisantes.

Afin de comprendre la biorthogonalité et les avantages qu'elle donne à la reconstitution (erreur de reconstitution très faible), se référer à l'annexe A2 où l'orthogonalité est explicitée avec les principes sur lesquels elle repose.

Partie C : Compression et reconstitution du signal ECG par ondelette

Les caractéristiques spécifiques des ondelettes en font un outil indispensable dans les applications d'engineering comme la compression, le dé-bruitage, le traitement d'image, etc... Dans ce qui suit, nous mettons en évidence les performances de cet outil dans le domaine de la compression du signal physiologique ECG.

3.5 La notion de compression

La compression en général est la manipulation d'un ensemble de données initiales par un codage spécifique et cela pour optimiser un espace de stockage. Elle est divisée en deux catégories : compression sans perte (lossless compression) ou compression avec perte (lossy compression) qui permet la reconstitution, [77,78].

La compression basée sur les transformées consiste à concentrer l'énergie du signal autour des coefficients de celle-ci. Dans le cas de la transformée en ondelettes l'information utile est cantonnée dans les basses fréquences [79].

3.5.1 Compression sans perte :

La compression est dite sans perte lorsqu'il n'y a aucune perte de données sur l'information d'origine après le cycle de compression / décompression. Il y a autant d'information après la compression qu'avant, elle est seulement réécrite d'une manière plus concise. Ces techniques sont utilisées pour de nombreux types de données notamment des documents, des fichiers exécutables ou des fichiers texte.

3.5.2 Compression avec perte :

Avec un algorithme de compression avec pertes, la suite de bits obtenue après les opérations de compression et de décompression est différente de l'originale, mais l'information reste sensiblement la même. Ce type de compression ne s'applique qu'aux données « perceptibles », en général sonores ou visuelles, qui peuvent subir une modification, parfois importante, sans que cela soit perceptible par un humain. Cependant, elle possède un fort taux de compression [80].

3.6 Principe de la compression

Dans la méthode de compression que nous allons utiliser, deux étapes importantes sont à prendre en considération : la décomposition et le seuillage.

En ce qui concerne la décomposition à bases d'ondelette les résultats précédents, nous ont conduits à utiliser les ondelettes bi orthogonales. Notons que le choix du niveau de la décomposition reste expérimental.

La compression consiste à ne conserver que les coefficients de détails d'ondelette les plus importants (en valeur absolue), les autres étant mis à zéro, puis on reconstitue de nouveau le signal à partir des coefficients d'ondelette seuillés. La figure 3.17 décrit la démarche de compression par seuillage des détails.

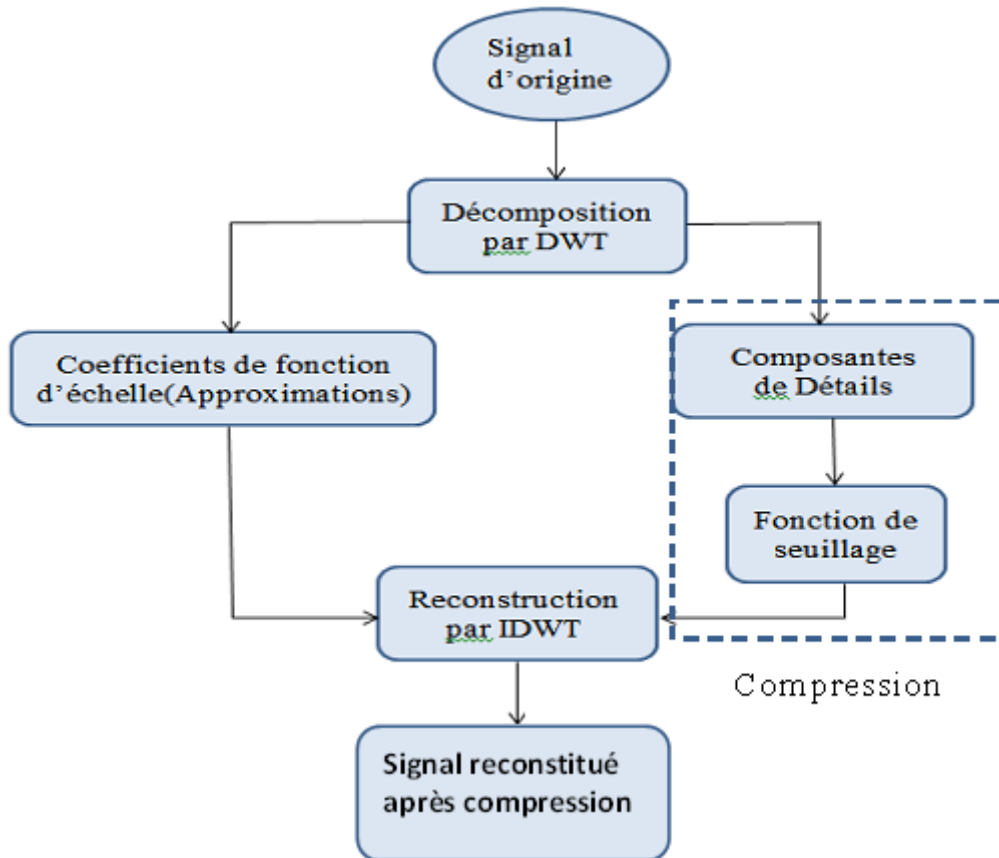


Figure 3.17 : Principe de compression-reconstitution.

3.6.1 Le seuillage

Cela revient à éliminer les détails les plus fins du signal. En effet, lors de la décomposition par ondelettes, plus la forme du signal analysé est en correspondance avec celle de l'ondelette analysante, plus le coefficient d'ondelette est important (le coefficient d'ondelette est issu de la corrélation entre l'ECG et l'ondelette).

De ce fait, si l'ondelette est correctement choisie, les coefficients d'ondelettes élevés correspondent aux zones contenant le plus d'informations. Ainsi en vue de compresser le signal on peut penser à annuler tous les coefficients qui sont en dessous d'un seuil.

3.6.1.1 Types de seuillage

a. Seuillage dur (hard thresholding)

Le seuillage dur consiste à fixer un seuil $S > 0$, à ne conserver que les coefficients d'ondelettes supérieurs à S et à mettre à zéro les autres :

$$\begin{cases} |x| < S \Rightarrow f(x) = 0 \\ |x| > S \Rightarrow f(x) = x \end{cases} \quad (3.11)$$

b. Seuillage doux (soft thresholding)

Dans le cas du seuillage doux, on met toujours à zéro les coefficients inférieurs à un seuil S . Par contre, pour ceux supérieurs à S , on atténue l'amplitude des coefficients par la valeur du seuil. Le coefficient seuillé sera donc plus petit que le coefficient du signal original.

$$\begin{cases} |x| < S \Rightarrow f(x) = 0 \\ |x| > S \Rightarrow f(x) = x - S \end{cases} \quad (3.12)$$

Les fonctions de seuillage des deux cas cités sont représentées par la figure 3.18, [81].

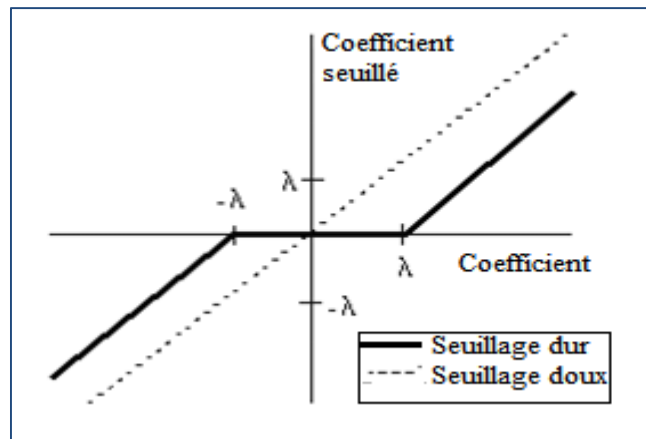


Figure3.18 : Fonctions de seuillage.

c. Seuillage statique

Cela revient à choisir un seul seuil qui sera fixe au cours de la compression pour tous les niveaux de décomposition. On l'appelle aussi seuillage global.

d. Seuillage dynamique

Ce genre de seuillage possède un seuil qui varie d'un niveau à un autre durant la compression de façon itérative. Notons que plus le nombre d'itérations est grand moins la compression sera importante mais par contre la décompression ou reconstitution sera bonne, c'est un compromis qu'il faudra prendre en compte par la suite [82].

Les précédentes définitions nous orientent vers le choix d'un seuillage statique dur pour une éventuelle compression. Car, en vue d'une reconstitution, il nous faut garder les valeurs des coefficients significatifs intacts.

Le seuillage dynamique doux est donc déconseillé, dans une compression mais plus adapté à un débruitage car il intervient en modifiant les valeurs des coefficients.

Notons que le choix du seuil et du niveau de décomposition N reste expérimental.

3.6.2 Application de la compression au signal ECG et reconstitution

Nous avons, à l'aide du logiciel MATLAB 7.1, réalisé et programmé un algorithme de compression avec reconstitution dans lequel nous avons introduit le traitement par ondelettes biorthogonales. Les étapes sont:

étape 1 : chargement et conversion des enregistrements issus de MIT-BIH Database (format 212, avec une extension .dat), en un format adapté au logiciel MATLAB 7.1 ;

étape 2 : extraction des deux dérivations contenues dans les différents enregistrements, (puisque nous avons voulu traiter chacune à part pour de meilleurs résultats) ;

étape 3 : choix de la valeur du seuil en fonction des valeurs des échantillons, pour chaque dérivation ;

étape 4 : compression à base d'ondelettes biorthogonales du signal ECG de chaque dérivation à plusieurs niveaux (cette compression consiste en une décomposition suivie d'un seuillage de type statique dur);

étape 5 : reconstitution du signal d'origine, basée elle aussi sur le traitement par ondelettes biorthogonales de chaque dérivation.

Nous allons par la suite exposer les différents résultats obtenus à l'aide de cet algorithme moyennant des estimateurs de taux de compression et de normes de reconstitution pour les différents signaux ECG de la MIT database.

3.7 Résultats pratiques et discussion

Les résultats obtenus sont représentés sous forme de tableaux comparatifs et de figures pour des signaux ECG normaux et d'autres révélant des pathologies, afin de montrer les performances de cette méthode dans les deux cas et de pouvoir ainsi élargir le domaine de son application.

Pour ce faire nous avons utilisé deux paramètres qui révèlent les performances obtenues grâce à l'application de notre algorithme :

- le paramètre de performance de compression (**PC**), qui représente le pourcentage des coefficients nuls après compression ainsi que le taux de compression. Il s'exprime par :

$$PC = \frac{\text{nbre.coeff.initial} - \text{nbre.coeff.final}}{\text{nbre.coeff.initial}} * 100 \quad (3.13)$$

$$= \frac{\text{nbre.coeff.nuls}}{\text{nbre.coeff.initial}} * 100,$$

la compression apporte un gain en volume de données ; la taille du signal est alors représentée par:

$$G = \frac{\text{nbre.coeff.final}}{\text{nbre.coeff.initial}} * 100 \quad (3.14)$$

$$= 100 - PC,$$

- celui de performance ou norme de reconstitution (**PR**) dont la formule est la suivante :

$$PR = \frac{100 * \|ECGr\|^2}{\|ECG\|^2} \quad (3.15)$$

avec : ECG le signal original et ECGr le signal reconstruit.

Nous avons étudié les différents paramètres utilisés pour obtenir des résultats optimaux de compression et reconstitution. Ces paramètres sont : le niveau de décomposition N, le seuil et l'ondelette biorthogonale la mieux adaptée. Pour cela, nous avons utilisé les ondelettes biorthogonales disponibles dans MATLAB7.1, en variant les niveaux de décomposition de N=1 à Nmax (niveau de décomposition maximum atteint par chaque ondelette). Par la suite, nous avons choisi un seuil qui prend les valeurs [5% ,10%, 15%, 20%] de la valeur maximale des échantillons du signal ECG pour chaque dérivation.

3.7.1 Etude du niveau de décomposition

Les figures 3.19, 3.20 sont des exemples de la compression avec reconstitution par ondelettes biorthogonales sur plusieurs ECG à deux dérivations.

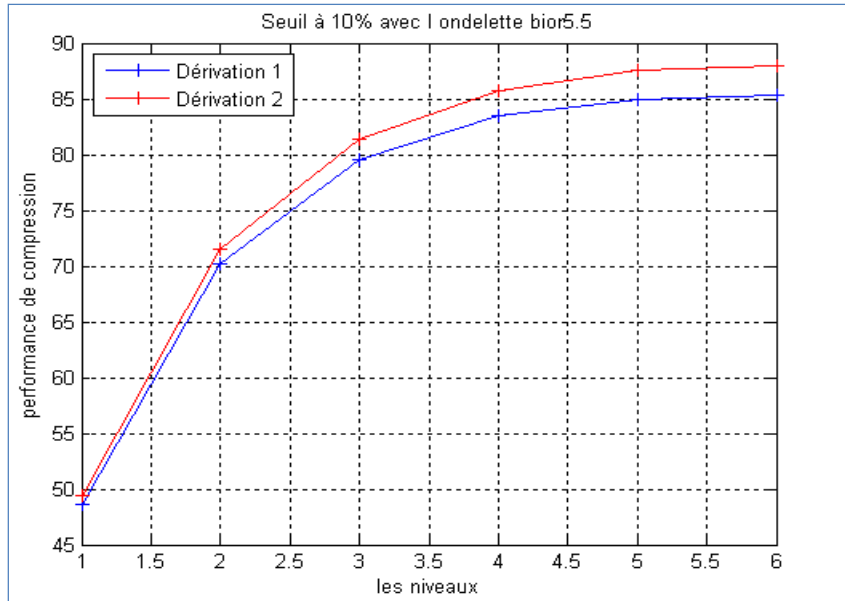


Figure 3.19: PC en fonction des niveaux de décomposition pour l'ECG 16273.

L'enregistrement 16273 est l'ECG d'un sujet sain ; nous avons appliqué l'ondelette « bior5.5 » avec un seuil de 10% à 8192 échantillons de ce dernier et nous avons varié le niveau de décomposition.

La figure 3.19 précédente montre le rapport entre le pourcentage de compression (PC) et les niveaux de décomposition pour les deux dérivation. Les résultats montrent un PC qui augmente avec ces niveaux et atteignant les 88%, ce qui est un très bon résultat. Donc si on s'arrête à cette comparaison, le niveau maximal est le bon choix ; mais il nous reste à voir l'erreur de reconstitution pour conclure.

Nous avons représenté la performance de reconstitution (PR) en fonction de ces mêmes niveaux de décomposition et pour les mêmes conditions (figure 3.20).

Le comportement dans ce cas est l'inverse. Le PR diminue si les niveaux de décomposition augmentent ; mais il n'est pas plus bas que 97%, ce qui est généralement acceptable.

Donc il existe un compromis qu'il faut prendre en considération dans le choix du niveau en fonction de l'application du programme.

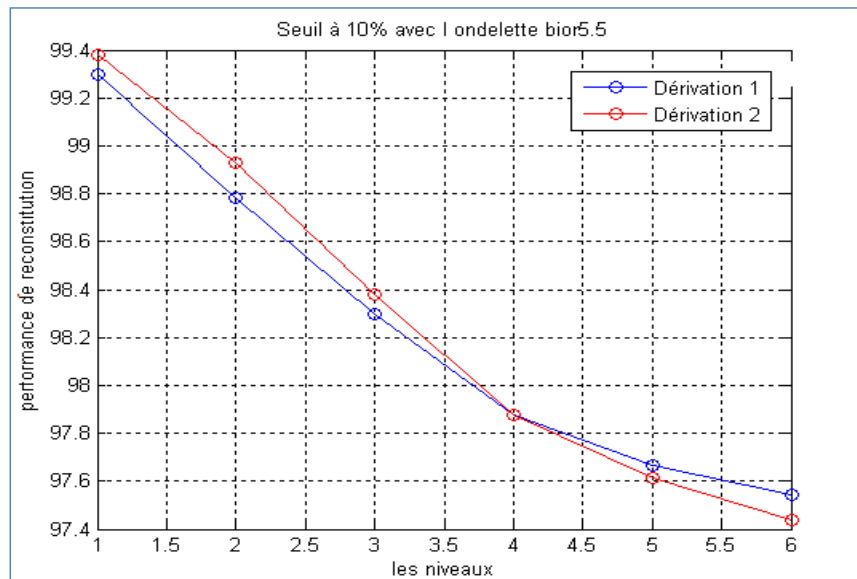


Figure 3.20 : PR en fonction des niveaux de décomposition pour l'ECG 16273.

En conclusion, l'étude pratique du comportement des PC et PR en fonction du niveau de décomposition montre que plus le niveau augmente plus la compression est importante, tandis que la reconstitution est moins bonne ; cette remarque est valable pour les deux dérivation des enregistrements utilisés et dans le cas de tous les ECG analysés (normaux et pathologiques), comme le montrent les figures ci-dessous.

Ce résultat est conforme à la théorie. Le fait d'appliquer plusieurs décompositions réduit de plus en plus le nombre de coefficients d'ondelette contenant l'information utile du signal. Ainsi le seuillage élimine plus de détails (bonne compression) et rend la reconstitution moins précise.

Or, étant donné que le signal ECG intervient dans le diagnostic médical, sa reconstitution doit être la plus fidèle possible. On doit donc favoriser la reconstitution au prix de la compression.

Prenons l'exemple de la figure 3.21 de l'ECG 103 qui est celui d'un sujet atteint d'Ectopie supraventriculaire (maladie qui altère la régularité du rythme cardiaque et qui est décelable à l'ECG).

Les figures 3.21 et 3.22 obtenues montrent les mêmes résultats que précédemment.

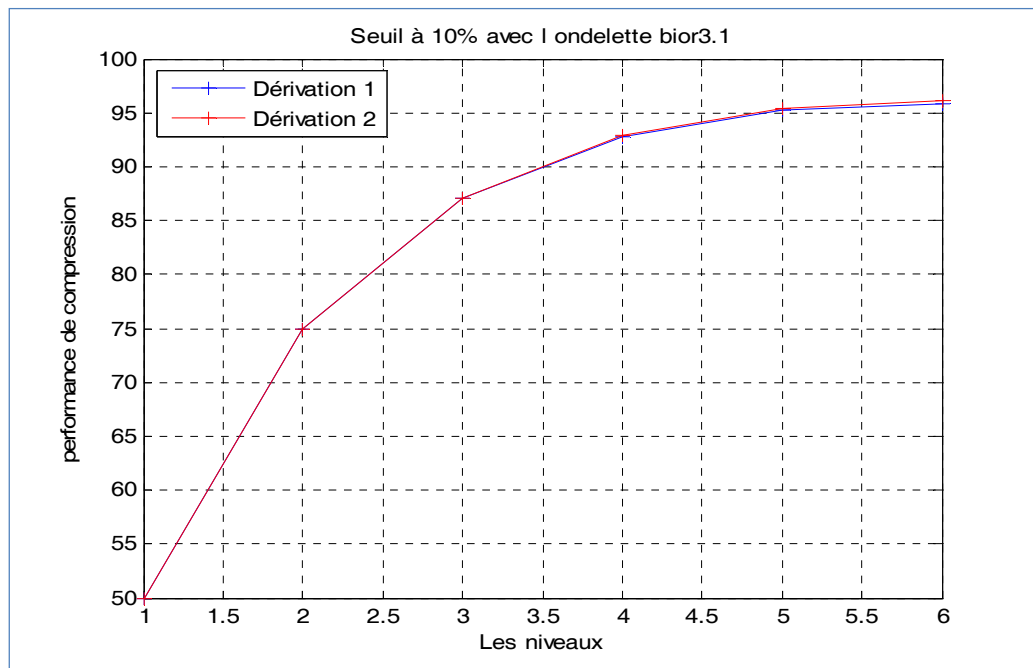


Figure 3.21 : PC en fonction des niveaux de décomposition pour l'ECG 103.

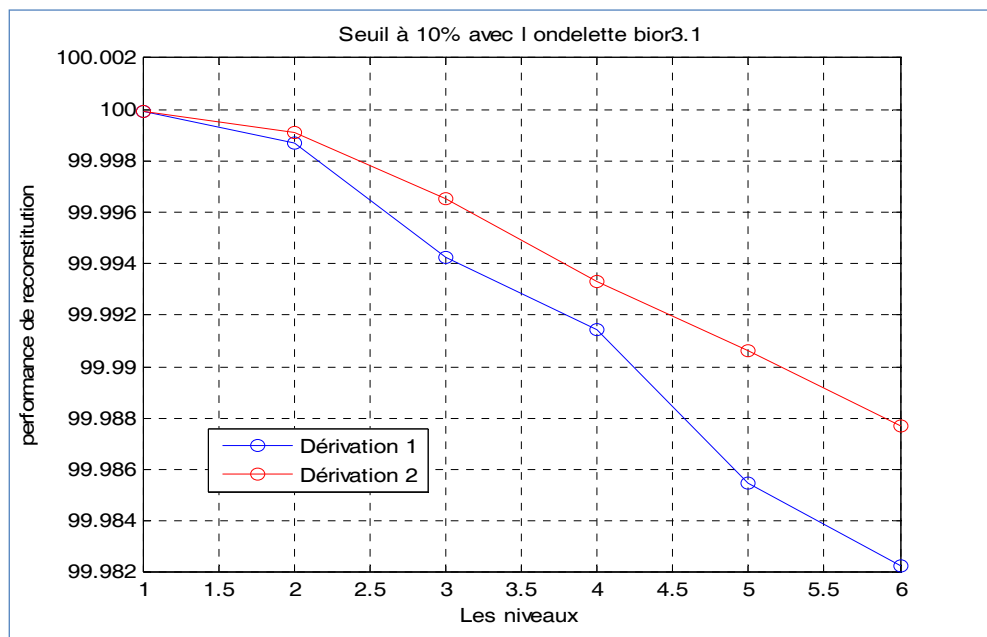


Figure 3.22 : PR en fonction des niveaux de décomposition pour l'ECG 103.

3.7.2 Etude de la valeur du seuil

Pour cette étude nous avons fixé le niveau de décomposition à 4 pour favoriser la reconstitution, comme nous l'avons déjà expliqué.

Nous avons appliqué l'algorithme de compression-reconstitution en variant les seuils pour tous les ECG dont nous disposons.

Prenons le cas de l'enregistrement cu01 qui est celui d'un sujet atteint de tachyarythmie ventriculaire, l'étude du comportement des PC et PR en fonction des seuils a fait apparaître que :

- dans le cas de la compression, plus le seuil augmente plus le PC est important ; il atteint les 90% pour les deux dérivation (figure 3.23),
- pour la reconstitution le PR diminue avec l'augmentation de la valeur du seuil (figure 3.24).

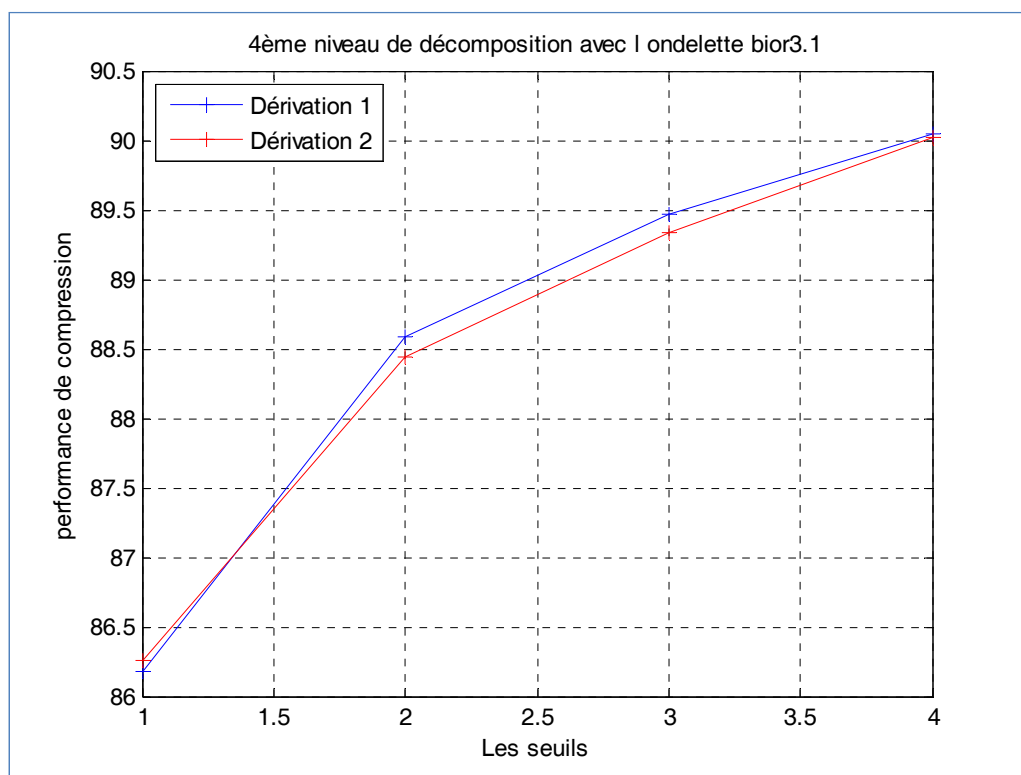


Figure 3.23 : PC en fonction des seuils pour l'ECG cu01.

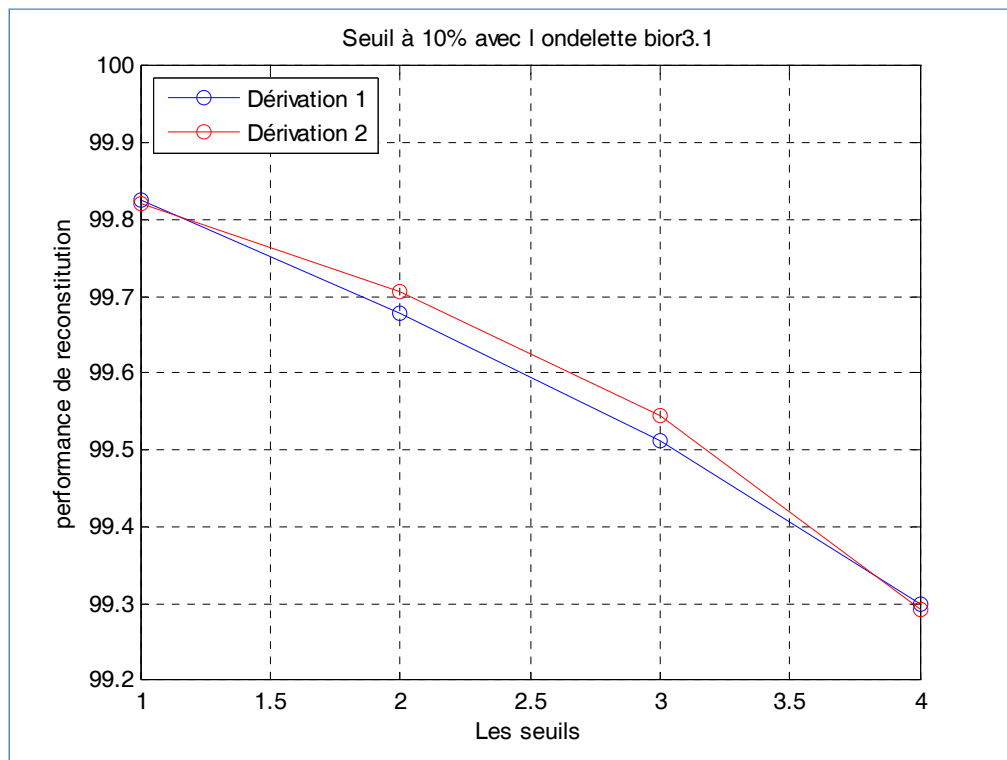


Figure 3.24: PR en fonction des seuils pour l'ECG cu01.

Ces résultats sont les mêmes pour tous les ECG analysés, ce qui n'est pas surprenant. Théoriquement si le seuil est grand on aura un nombre important d'échantillons nuls après seuillage, ce qui induit une bonne compression ; mais ces détails éliminés altéreront le signal lors de sa reconstitution.

Notons que dans le cas de la variation des niveaux (figures 3.20, 3.22) et celle des seuils, les PR sont toujours très bons. Même s'ils diminuent, ce qui est dû à l'ondelette biorthogonale qui offre par ses caractéristiques, une reconstitution optimale du signal ECG.

Donc, tant que la compression avec reconstitution est à base d'ondelettes biorthogonales, quels que soient le niveau et le seuil choisis, la performance de reconstitution est satisfaisante ; mais puisque le signal analysé est un signal médical, on doit favoriser la reconstitution et soumettre le signal obtenu à l'avis d'un médecin pour tirer des conclusions finales sur le choix du niveau et du seuil.

Les figures 3.25 et 3.26 représentent un autre exemple dans l'étude de la valeur du seuil pour un sujet sain (l'enregistrement X_16265) :

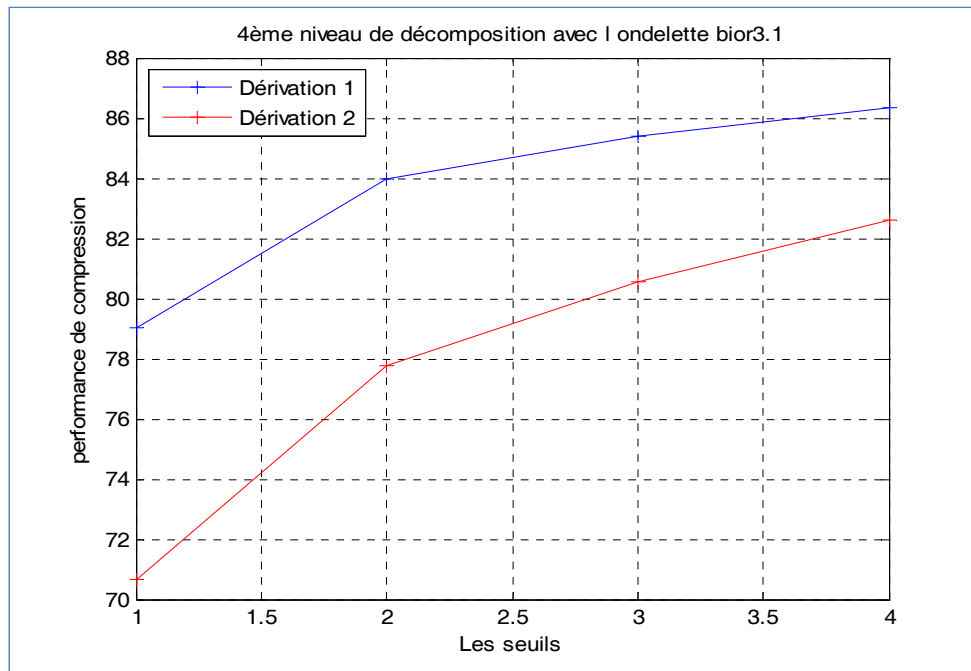


Figure 3.25 : PC en fonction des seuils pour l'ECG X_16265.

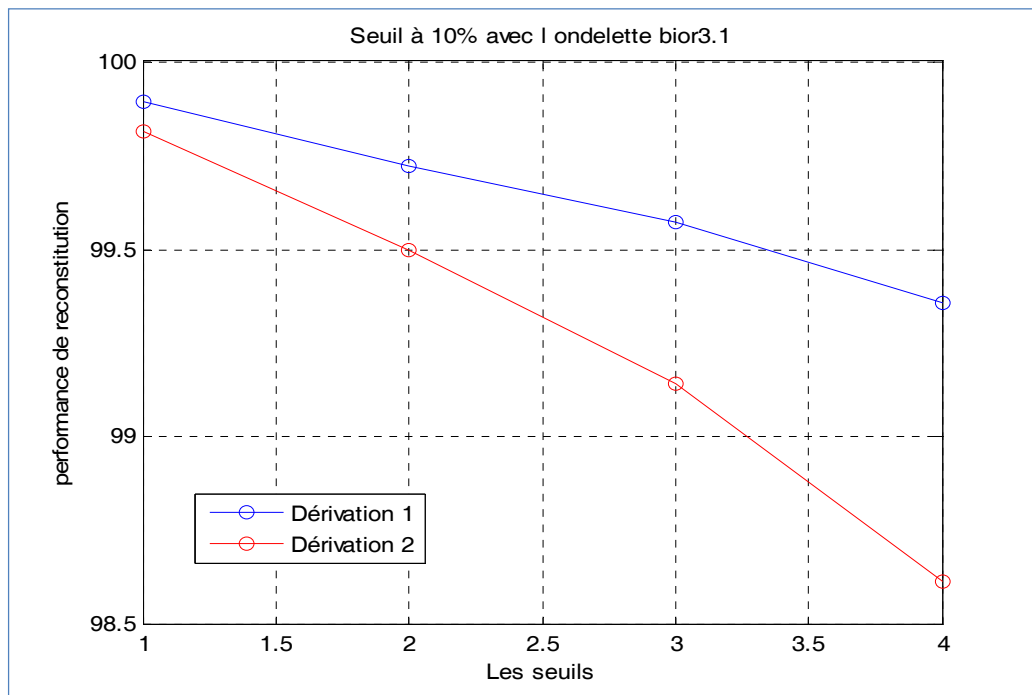


Figure 3.26 : PR en fonction des seuils pour l'ECG X_16265.

3.8 Tests et résultats

Suite à l'étude du niveau de décomposition et de celle du seuil, nous avons appliqué un algorithme de compression à chaque enregistrement dont nous disposons sain et pathologique. Les tableaux suivants montrent les meilleurs résultats obtenus dans chaque cas.

1^{er} Cas : les ECG analysés représentent des sujets sains ; le tableau 3.5 montre l'exemple de l'enregistrement 16273.

Tableau 3.5 : Résultats pour l'ECG 16273

Ondelette	Niveau	Seuil (%)	Première dérivation			Deuxième dérivation		
			PC (%)	G (%)	PR (%)	PC (%)	G (%)	PR (%)
bior3.1	6	10	84.6557	15,3443	99.7875	85.4235	14,5765	99.5904

Nous avons appliqué ce programme sur 8192 échantillons de l'ECG 16273, nous avons atteint un pourcentage de 85% de coefficients nuls, ce qui est un résultat satisfaisant.

Nous avons montré l'ECG reconstitué de la figure 3.27(b) à un cardiologue. Ce dernier a pu l'interpréter sans difficulté.

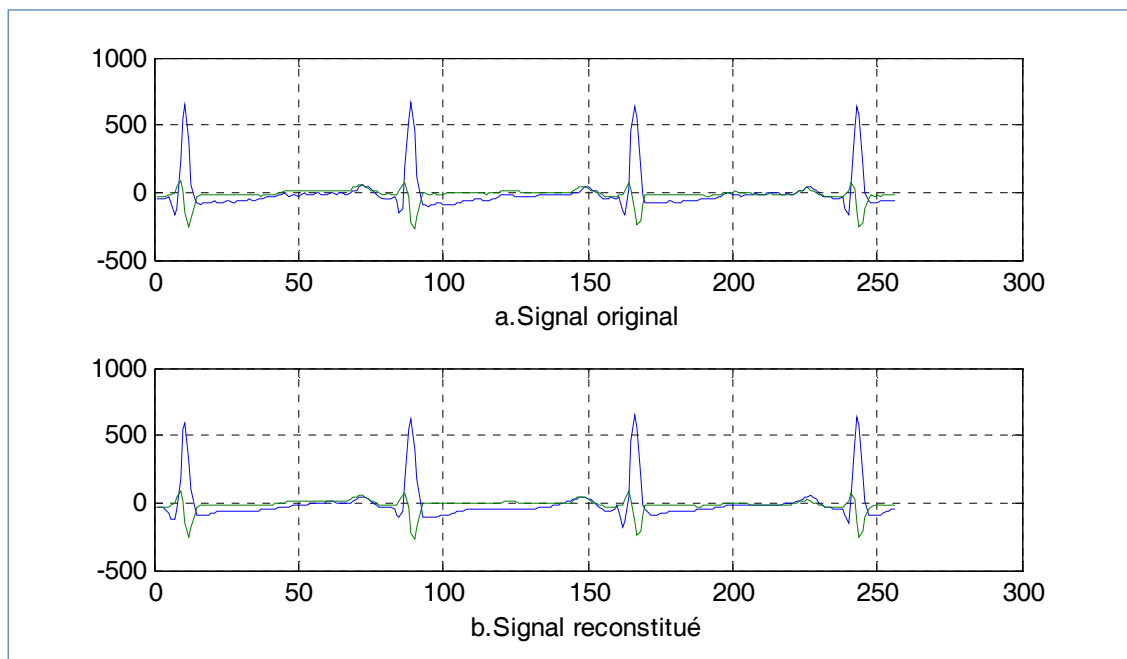


Figure 3.27 : Signal original et signal reconstitué ECG 16273.

2^{ème} Cas : les ECG analysés représentent des sujets atteints de tachyarythmie ventriculaire; (Tableau 3.6).

Tableau 3.6 : Résultats pour l'ECG cu01

Ondelette	Niveau	Seuil (%)	Première dérivation			Deuxième dérivation		
			PC (%)	G (%)	PR (%)	PC (%)	G (%)	PR (%)
bior3.1	6	10	89.8111	10,1889	99.7944	89.7745	10,2255	99.8000

Nous avons atteint dans ce cas les 89.8% de compression. La reconstitution soumise à l'avis du médecin a été approuvée, la figure 3.28 en montre l'allure.

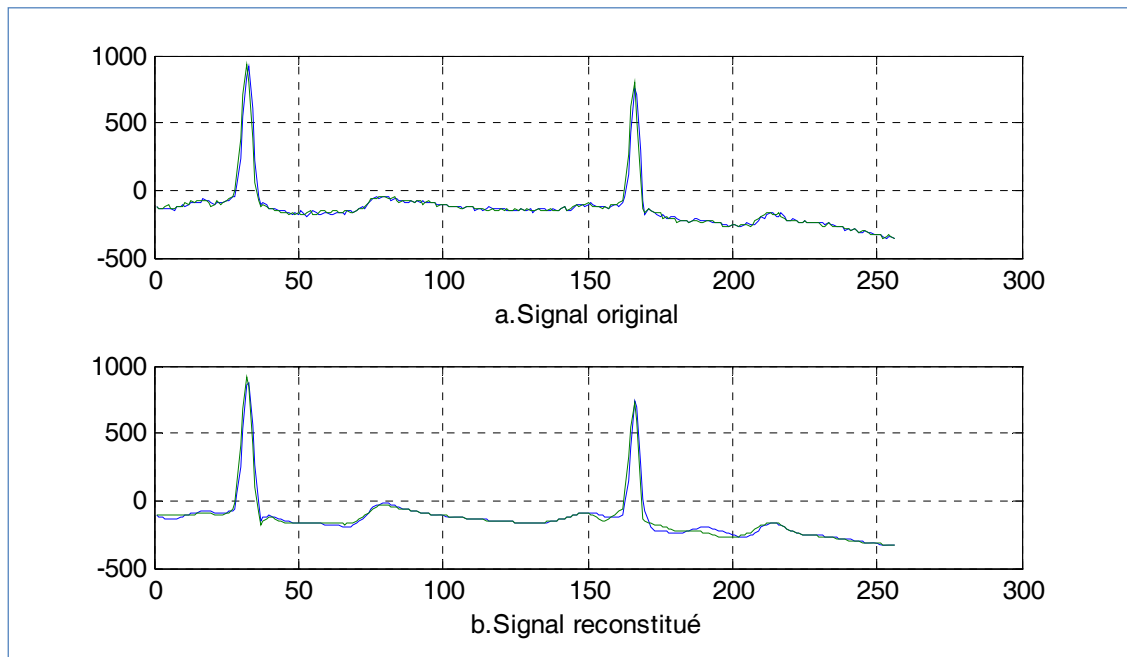


Figure 3.28 : Signal original et signal reconstitué ECG cu01.

3^{ème} Cas : les ECG analysés représentent des sujets atteints de fibrillation atriale ; le tableau 3.7 montre l'exemple de l'enregistrement X_04936

Tableau 3.7 : Résultats pour l'ECG X_04936

Ondelette	Niveau	Seuil (%)	Première dérivation			Deuxième dérivation		
			PC (%)	G (%)	PR (%)	PC (%)	G (%)	PR (%)
bior3.3	6	10	76.0510	23,949	99.9578	89.5261	10,4739	99.8267

Les résultats obtenus dans ce 3^{ème} cas sont tout aussi satisfaisants que dans les exemples précédents.

Comme le montre la figure 3.29, le signal reconstitué présente les mêmes caractéristiques que le signal original et permet ainsi un diagnostic identique.

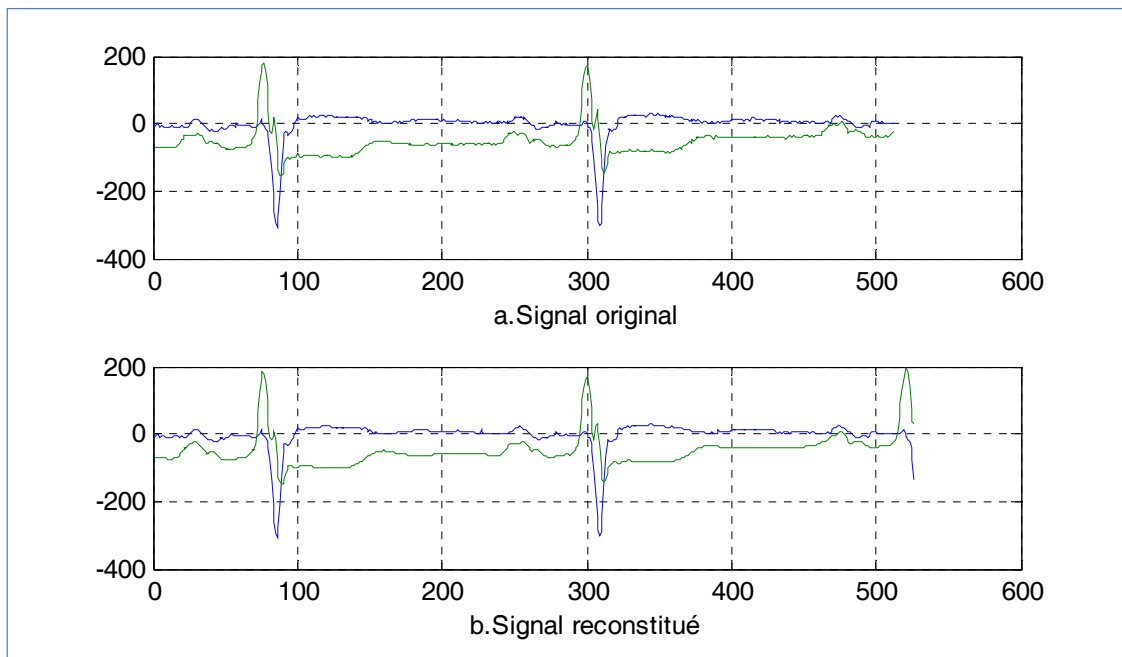


Figure 3.29 : Signal original et signal reconstitué ECG X_04936.

4^{ème} Cas : les ECG analysés représentent des sujets atteints d'ectopie ventriculaire ; le tableau 3.8 montre l'exemple de l'enregistrement 105 . Les PC pour les deux dérivation sont très élevés ainsi que les PR, ce qui montre que la méthode est performante dans les deux cas, figure 3.30.

Tableau 3.8 : Résultats pour l'ECG 105.

Ondelette	Niveau	Seuil (%)	Première dérivation			Deuxième dérivation		
			PC (%)	G (%)	PR (%)	PC (%)	G (%)	PR (%)
bior4.4	3	5	87.2322	12,7678	99.9978	87.4513	12,5487	99.9988

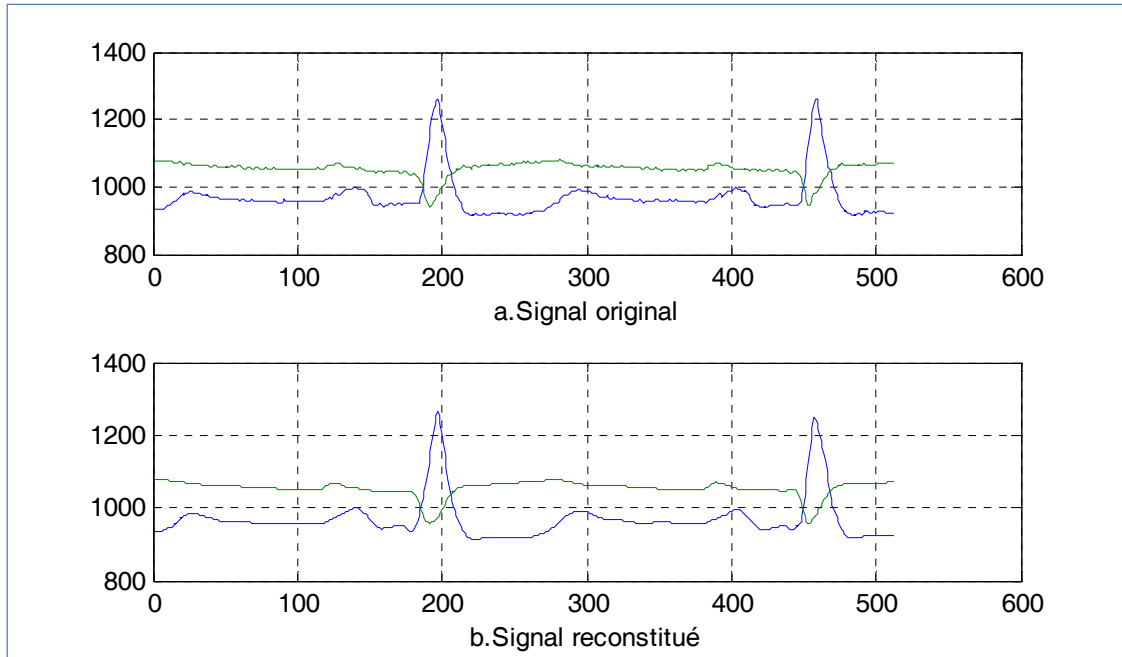


Figure 3.30 : Signal original et signal reconstitué ECG 105

La synthèse des différents résultats obtenus par cette méthode de compression nous conduit vers les conclusions suivantes.

- Nous avons pu démontrer que l'ondelette biorthogonale est la mieux adaptée à la compression avec reconstitution du signal ECG ; donc quelle que soit l'ondelette biorthogonale choisie, nous pouvons nous attendre à de bons taux de compression et une assez bonne reconstitution.
- Une étude expérimentale poussée nous a permis de choisir l'ondelette biorthogonale la mieux adaptée aux formes des ECG sains et pathologiques.

Nos résultats ont fait l'objet de 2 communications [83,84].

Conclusion

La DWT biorthogonale a montré des performances de compression intéressantes et a permis une reconstitution fidèle de nos signaux. Ainsi, le diagnostic des diverses pathologies cardiaques peut à présent se fonder sur la représentation non altérée des données ECG transmises et reconstituées. Les résultats obtenus présentent des taux de compression satisfaisants (de 80% à 90%) et des reconstitutions approuvées par des cardiologues.

Introduction

Afin de pouvoir diagnostiquer une anomalie ou pathologie cardiovasculaire, il est nécessaire d'acquérir le signal physiologique qui est de nature électrique (l'ECG), de le traiter et enfin l'analyser. L'analyse permettra en conséquence la prise de décision relative à la présence ou non d'une anomalie. La décision dépendra du procédé d'analyse basé sur des méthodes en adéquation avec les spécificités du signal. En effet la non stationnarité du signal cardiaque ainsi que la non linéarité des phénomènes qui le génèrent, imposent des méthodes d'analyse en mesure d'extraire efficacement les informations pertinentes.

Dans nos travaux de recherche, nous avons proposé deux approches intéressantes pour l'extraction des paramètres qui sont nécessaires pour la classification et donc une bonne discrimination entre les signaux ECG normaux et anormaux. Ces approches ont fait l'objet de deux publications internationales [85,86].

La première approche utilise une méthode temporelle basée sur la modélisation du signal ECG par les réseaux de neurones RBF (réseaux à bases radiales) combinés à un algorithme d'apprentissage automatique appelé Generalized Orthogonal Forward Regression (GOFR) et une classification supervisée moyennant l'algorithme Support Vecteurs Machine SVM. Cette approche permet de concevoir un système qui fonctionne en temps réel à partir d'une extraction en ligne des informations et donc d'établir une aide au diagnostic médical.

La seconde approche est une méthode temps-fréquence qui repose sur la transformée en ondelettes discrète DWT dont la décomposition en bandes de fréquences recouvre les caractéristiques fréquentielles des différentes ondes de l'ECG (P, QRS, T). L'exploitation des cinq premiers niveaux de décomposition a permis l'extraction des coefficients de détails pertinents de chaque signal ECG. La variance de l'énergie (fonction statistique) de ces coefficients a été calculée pour représenter leur distribution. Les méthodes de réduction de paramètres ACP ou ICA ont été combinées au même classifieur SVM.

Notre objectif principal, consiste à démarquer les signaux normaux des signaux pathologiques et ce, moyennant les deux techniques d'extraction de paramètres. Les comparaisons seront fondées sur deux mesures de la performance issues des matrices de confusion: la spécificité et la sensibilité.

Avant de présenter les différentes étapes de tests avec interprétation des résultats des deux approches sur les signaux réels, nous allons résumer la théorie des outils mathématiques

utilisés, à savoir : l'analyse en composante principale, indépendante ainsi que la classification supervisée par SVM.

4.1 Analyse en composante principale (ACP)

Le but de l'analyse en composantes principales (ACP) est de condenser les données originales déduites de l'analyse des différentes ondes de l'ECG après décomposition par la DWT, en de nouveaux groupements de façon à ce qu'ils ne présentent pas de corrélation entre eux et soient ordonnés en terme de pourcentage de variance (énergie) apportée par chaque composante.

Le recours à cette approche relevant de: l'analyse en composantes principales, vise des traitements numériques tels que la classification automatique, [87-90].

L'ACP est une méthode qui consiste à transformer des variables "corrélées" en statistique en nouvelles variables indépendantes nommées "composantes principales", ou axes. Il s'agit d'une approche à la fois géométrique (représentation des variables dans un nouvel espace géométrique selon des directions d'inertie maximale) et statistique (recherche d'axes indépendants expliquant au mieux la variance des données). Lorsqu'on veut alors compresser un ensemble de N variables aléatoires, les n premiers axes de l'ACP sont un meilleur choix, du point de vue de l'inertie ou de la variance, [90].

On applique usuellement une ACP sur un ensemble de p variables aléatoires $E_1, \dots, E_p$ connues à partir d'un échantillon de n réalisations conjointes de ces variables. Dans notre étude, ces variables sont les valeurs d'énergie calculées à partir des coefficients de détail issus de la DWT. n est le nombre total des signaux normaux et pathologiques. L'échantillon de ces p valeurs d'énergie, peut être structuré dans une matrice E à n lignes et p colonnes :

$$E = \begin{bmatrix} E_{1,1} & \cdot & E_{1,p} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ E_{n,1} & \cdot & E_{n,p} \end{bmatrix} \quad (4.1)$$

p = nombre de niveaux de décomposition.

n = nombre de signaux.

4.1.1. Transformations de l'échantillon –Matrice normalisée

Le vecteur $(\bar{M}_1, \dots, \bar{M}_p)$ est la moyenne de chaque colonne de la matrice E qu'on note $\bar{M}$ tel que :

$$\bar{M}_j = \frac{1}{n} \sum_{i=1}^n E_{i,j} \quad (4.2)$$

$$\bar{M}_j = [\bar{M}_1 \quad . \quad . \quad . \quad \bar{M}_p]$$

On appelle donnée centrée le tableau de données :

$$\bar{E} = (E_{i,j} - \bar{M}_j)_{i=1,\dots,n, j=1,\dots,p} \quad (4.3)$$

L'application de la relation (4.2) donne la matrice centrée $\bar{E}$.

$$\bar{E} = \begin{bmatrix} E_{1,1} - \bar{M}_1 & . & . & . & E_{1,p} - \bar{M}_p \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ E_{n,1} - \bar{M}_1 & . & . & . & E_{n,p} - \bar{M}_p \end{bmatrix} \quad (4.4)$$

Les écarts types sont donnés par la relation (4.5).

Pour tout $j=1,\dots,p$, nous avons :

$$\sigma_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (E_{i,j} - \bar{M}_j)^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{E}_{i,j})^2} \quad (4.5)$$

Et la variance de l'énergie est donnée par :

$$\text{Var} = \frac{1}{n} \sum_{i=1}^n (E_{i,j} - \bar{M}_j)^2 \quad (4.6)$$

Les données centrées et réduites sont données alors par :

$$\tilde{E} = \left(\frac{E_{i,j} - \bar{M}_j}{\sigma_j} \right)_{i=1,\dots,n, j=1,\dots,p} = \left(\frac{\bar{E}_{i,j}}{\sigma_j} \right)_{i=1,\dots,n, j=1,\dots,p} \quad (4.7)$$

L'application de cette relation (4.7) donne la matrice centrée et réduite :

$$\tilde{E} = \begin{bmatrix} \frac{E_{1,1}-\bar{M}_1}{\sigma_1} & \dots & \dots & \frac{E_{1,p}-\bar{M}_p}{\sigma_p} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{E_{n,1}-\bar{M}_1}{\sigma_1} & \dots & \dots & \frac{E_{n,p}-\bar{M}_p}{\sigma_p} \end{bmatrix} \quad (4.8)$$

4.1.2. Calcul de covariance et de la corrélation

Une fois la matrice E transformée en $\bar{E}$ ou $\tilde{E}$, on calcule la matrice de covariance $\text{Cov}(E)$ et de corrélation $\text{Cor}(E)$, il suffit de la multiplier par sa transposée .

- la matrice de variance-covariance des $E_1, \dots, E_n$ si E est la matrice centrée est donnée par: $\bar{M}_j^T$

$$\text{Cov}(E) = \left(\frac{1}{n} \sum_{i=1}^n (E_{i,j} - \bar{M}_j) (E_{i,j}^T - \bar{M}_j^T) \right)_{i=1, \dots, n, j=1, \dots, p} = \frac{1}{n} \bar{E}^T \bar{E} \quad (4.9)$$

- la matrice de corrélation des $E_1, \dots, E_n$ si E est centrée et réduite est donnée par :

$$\text{Cor}(E) = \left(\frac{1}{n} \sum_{i=1}^n \frac{(E_{i,j} - \bar{M}_j)(E_{i,j}^T - \bar{M}_j^T)}{\sigma_i \sigma_j} \right)_{i=1, \dots, n, j=1, \dots, p} = \frac{1}{n} \tilde{E}^T \tilde{E} \quad (4.10)$$

Ce sont des matrices carrées (de taille p), symétriques et réelles. Elles sont donc diagonalisables dans une base orthonormée.

Nous considérerons que le nuage est transformé (centré et réduit si besoin est). Chaque $E_{i,j}$ est donc remplacé par $E_{i,j} - \bar{M}_j$ ou $\frac{E_{i,j} - \bar{M}_j}{\sigma_j}$.

Nous notons $\bar{E}$, la matrice E si $E_{i,j}$ est remplacé par $E_{i,j} - \bar{M}_j$ ou $\tilde{E}$ la matrice E si $E_{i,j}$ est remplacé par $\frac{E_{i,j} - \bar{M}_j}{\sigma_j}$.

Le principe de l'ACP est de trouver un axe u , issu d'une combinaison linéaire des $E_{i,j}$, telle que la variance du nuage autour de cet axe soit maximale. Pour bien comprendre, imaginons que la variance de u soit égale à la variance du nuage; on aurait alors trouvé une combinaison des $E_{i,j}$ qui contient toute la diversité du nuage original (toute la part de sa diversité captée par la variance).

4.1.3 Projection :

Finalement, nous cherchons le vecteur u tel que la projection du nuage sur u ait une variance maximale. La projection de l'échantillon des E sur u s'écrit :

$$P_U(E) = E u \quad (4.11)$$

la variance empirique de $P_U(E)$ vaut donc :

$$\begin{aligned} (P_U(E))^T \frac{1}{K} (P_U(E)) &= u^T E^T \frac{1}{K} E u \\ C &= E^T \frac{1}{K} E \end{aligned} \quad (4.12)$$

Où C est la matrice de covariance.

Comme nous avons vu plus haut que C est diagonalisable dans une base orthonormée, notons X le changement de base associé et Δ la matrice diagonale formée de son spectre :

$$(P_U(E))^T \frac{1}{K} (P_U(E)) = Xu^T \Delta Xu \quad (4.13)$$

Après cette réécriture, nous cherchons le vecteur unitaire Xu qui maximise $Xu^T \Delta Xu$,

Où $\Delta = \text{Diag}(\lambda_1, \dots, \lambda_p)$ est diagonale (les valeurs de la diagonale de Δ sont rangées par ordre décroissant). On peut rapidement vérifier qu'il suffit de prendre le premier vecteur unitaire. Nous avons alors :

$$Xu^T \Delta Xu \quad (4.14)$$

Nous obtenons ainsi le résultat suivant:

- u est vecteur propre de C associé à la valeur propre λ_1
- u est de norme 1.

La valeur propre λ_1 est la variance empirique sur le premier axe de l'ACP. On continue la recherche du deuxième axe de projection w , avec le même principe, en posant la condition d'orthogonalité par rapport à u .

4.1.4. Diagonalisation

La diagonalisation de la matrice de covariance (ou de corrélation si on se place dans un modèle

réduit), nous a permis d'écrire que le vecteur qui explique le plus d'inertie du nuage est le premier vecteur propre. De même le deuxième vecteur qui explique la plus grande part de l'inertie restante est le deuxième vecteur propre, etc. Nous avons vu en outre que la variance expliquée par le p -ième vecteur propre vaut λ_p . Finalement, la question de l'ACP se ramène à un problème de diagonalisation de la matrice de corrélation.

4.1.5.Choix de la dimension de projection:

On appelle pourcentage d'inertie reproduite en dimension k la valeur P :

$$P = \frac{\sum_{j=1}^k \lambda_j}{\sum_{j=1}^p \lambda_j} \quad (4.15)$$

Le choix des axes de projection dépend de la valeur de cette grandeur, [91-93] .

En pratique : utiliser $\text{cor}(x)$ lorsque les variables n'ont pas les mêmes ordres de grandeurs (unités différentes, par exemple). Dans la suite, on notera C la matrice choisie. L'ACP n'est pas une méthode de classification ; Elle permet la réduction de dimensions seulement.

4.2 Analyse en Composantes Indépendantes (ACI)

L'analyse en composantes indépendantes (ICA) est en réalité une extension d'une méthode d'analyse de données en composantes principales (ACP). En fait, si l'ACP cherche à extraire des variables décorrélées, en se limitant à imposer la contrainte d'indépendance aux statistiques d'ordre deux (matrice de covariance), l'ACI en revanche, cherche l'indépendance des statistiques d'ordre supérieur à deux des variables mesurées. C'est une méthode statistique récemment développée qui traite des observations vectorielles (multivariées) dans le but de trouver une représentation linéaire des données non gaussiennes afin que les composantes soient statistiquement indépendantes. Cette idée simple s'est révélée très fructueuse pour le traitement des signaux et des images multi-capteurs dans de nombreux domaines tels que : les interférences des ondes radio, les réseaux d'antennes pour les télécommunications, la prise de son, les signaux biomédicaux multi-électrodes, l'imagerie biomédicale, les images de télédétection (optiques et radar) [94,95].

L'analyse en composantes indépendantes d'un vecteur aléatoire réel x de taille p consiste à estimer le modèle génératif des données :

$$X = A S \quad (4.16)$$

Où: A : est la matrice de mélange.

S : est un vecteur aléatoire de taille n .

Le but est donc de retrouver la matrice de séparation W telle que :

$$S = W X \quad (4.17)$$

de telle manière que les composantes de S soient aussi indépendantes que possible au sens de la maximisation d'une fonction $F(S)$ qui est une mesure de l'indépendance statistique[12].

Les méthodes d'ACI reposent donc, sur une mesure d'indépendance qui reste à définir. Les différentes approches se différencient par la mesure d'indépendance retenue et par la méthode d'optimisation choisie.

Les principaux algorithmes[94,97,98] connus dans la littérature, qui permettent de présenter un panel varié d'approches et de résoudre le problème d'ACI ont été énoncés en Annexe A3.

Dépendant de la nature du problème de classification (vecteur de données, choix des caractéristiques, et choix du classificateur), la réduction de dimensions peut être fondamentale pour avoir de bonnes performances de classification.

4.3 Classification

La classification permet de différencier entre deux ou plusieurs classes d'état de santé et donc constitue un principe de base dans l'aide au diagnostic.

Dans nos travaux, nous avons proposé deux approches de traitement des signaux ECG pour atteindre le meilleur taux de classification et donc une bonne discrimination entre les signaux ECG normaux et anormaux. Dans la littérature, plusieurs méthodes de calcul des fonctions discriminantes permettent d'estimer la classe d'un nouvel exemple. L'ensemble des signaux ECG que nous utilisons pour l'apprentissage est labélisé. Ce qui nous oriente vers un contexte de classification supervisée. Les méthodes couramment utilisées en génie-biomédical, sont: les réseaux de neurones artificiels (ANN: Artificial Neural Network) [99-102], les K-plus proches voisins (K-NN: K-Nearest Neighbors) [103,104], les réseaux de neurones à fonction de

base radiale (RBFNN: Radial Basis Function Neural Network) [105], les réseaux de neurones récurrents (RNN: Recurrent Neural Network) [106,107] et les machines à vecteurs supports (SVM: Support Vector Machine) [108].

Nous avons utilisé la méthode de classification : SVM. Nous avons retenu la méthode des machines à vecteurs supports SVM car elle présente différents avantages. Cette méthode ne fait aucune hypothèse sur la forme des classes, fonctionne avec de faibles populations d'apprentissage, possède une excellente capacité de généralisation, permet de s'affranchir des problèmes des espaces de représentation de trop grande dimension, évite le sur-apprentissage et possède une solution unique.

4.3.1 Classification par machines à vecteurs support

La classification SVM est une méthode de classification supervisée binaire (à deux classes). Elle fut introduite par Vapnik en 1995. Le principe des SVM consiste à définir, par apprentissage, une surface de séparation entre les attributs (ou point) des classes positives et négatives minimisant le risque d'erreur et maximisant la marge entre les classes [109-112].

Soient les données d'entrée $x = (x_1 \dots x_N)$ associées aux sorties du classificateur (étiquettes) $y = (y_1 \dots y_N)$ dont les valeurs sont $\{+1, -1\}$ (les deux classes). Soit H , l'hyperplan séparateur (frontière de décision) qui a pour équation :

$$w \cdot x + b = 0 \quad (4.18)$$

L'hyperplan optimal est celui pour lequel la distance (marge) aux points les plus proches (vecteurs supports) est maximale (Figure 4.1). De plus, H doit vérifier les conditions suivantes :

$$\begin{cases} w \cdot x_i + b \geq +1 \\ w \cdot x_i + b \leq -1 \end{cases} \quad (4.19)$$

Ce qui est équivalent à : $y_i(w \cdot x_i + b) \geq 1$ pour $i = 1, \dots, N$

La marge géométrique, qui est la distance entre deux classes vaut $2/\|w\|$. Maximiser cette marge revient à minimiser son inverse. Le classificateur SVM est la solution d'un problème d'optimisation quadratique

$$\begin{cases} \text{minimiseer : } \frac{\|w\|^2}{2} \\ \text{sous la contrainte : } y_i(w \cdot x_i + b) \geq 1 \quad \text{pour } i = 1, \dots, N \end{cases} \quad (4.20)$$

La fonction de discrimination associée à l'hyper plan optimal devient

$$f(x) = \sum_{i \in S} \alpha_i y_i < x_i, x > + b \quad (4.21)$$

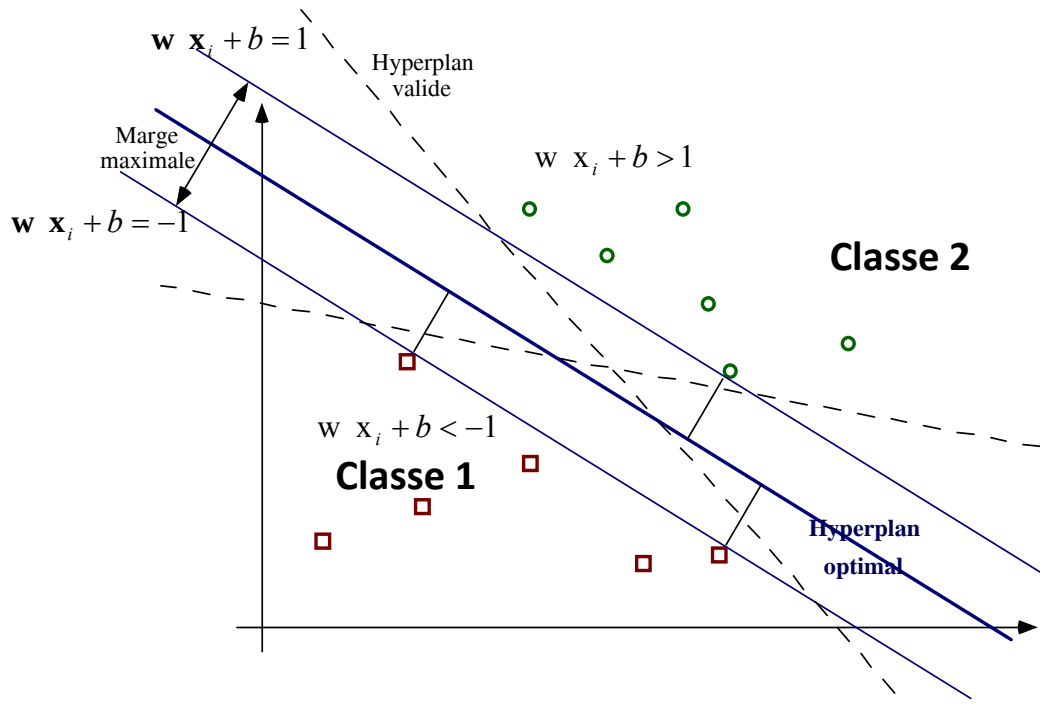


Figure 4.1 Principe des SVM : Hyperplan maximisant la marge.

Pour améliorer les performances de classification présentés ci-dessus il faut faire appel aux SVM non linéaires. Cette méthode repose sur l'existence d'un classificateur linéaire dans un espace approprié de dimension supérieure, c'est pourquoi des fonctions dites noyaux (ou kernel) sont utilisées pour une séparation optimale des données. Le choix du noyau se fait à priori de l'influence sur les résultats de classification [113].

Notons « Ω » une transformation non linéaire de l'espace d'entrée « X » en un espace de re-description « $\Omega(X)$ » de dimension supérieure. Alors, la formulation faite pour l'espace « X » doit se faire pour l'espace « $\Omega(X)$ », en notant :

$$K(x_i, x_j) = \langle \Omega(x_i), \Omega(x_j) \rangle \quad (4.22)$$

La fonction de discrimination s'écrit : $f(x) = \sum_{i \in S} \alpha_i y_i K(x_i, x_j) + b \quad (4.23)$

La figure 4.2 montre que les données non linéairement séparables dans l'espace d'entrée « X » deviennent séparables dans l'espace augmenté $\Omega(X)$.

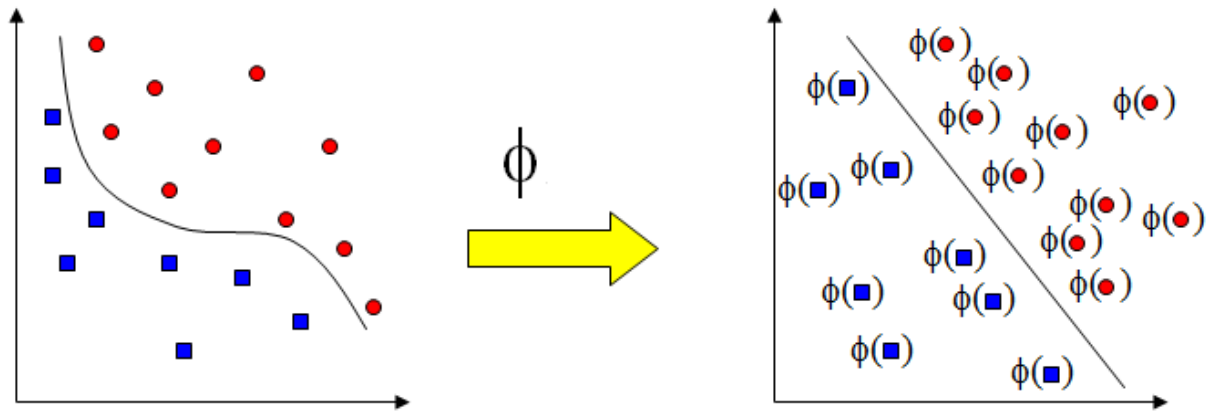


Figure 4.2 : Illustration de l'effet de changement d'espace par une fonction noyau.

Afin d'évaluer notre méthode de classification, nous définissons la matrice de confusion également appelée tableau de contingence. Cette matrice est représentée par le tableau 4.1 [114], où chaque colonne représente le nombre d'occurrences d'une classe estimée, tandis que chaque ligne représente le nombre d'occurrences d'une classe réelle (ou de référence). 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 de signaux ECG de la classe considérée.

Tableau 4.1 : Matrice de confusion

		Classes estimées	
		Cas Normal	Cas Anormal
Classes réelles	Cas Normal	TP	FN
	Cas Anormal	FP	TN

Où TP désigne les vrais positifs (True Positive), c'est le nombre de battements normaux classés normaux,

TN les vrais négatifs (True Negative), c'est le nombre de battements anormaux classés anormaux

,
FP désigne les faux positifs (False Positive), c'est le nombre de battements anormaux classés normaux

et FN les faux négatifs (False Negative), c'est le nombre de battements normaux classés anormaux.

Evaluer les performances d'un système de classification est un enjeu de grande importance car ces performances peuvent être utilisées pour l'apprentissage en tant que tel ou pour optimiser les valeurs des hyper paramètres du classificateur. La première mesure à laquelle nous allons nous intéresser, est le taux de bonne classification ou la précision. Il s'agit de l'indicateur le plus évident permettant d'évaluer les performances d'un système de classification.

a. **Le Taux de classification (Précision) :**

Nous définissons le taux de classification ou Accuracy comme suit :

$$T_c = Acc = \frac{\text{nombre d'échantillons bien classés}}{\text{nombre total d'échantillons}} * 100\%$$

$$T_c = Acc = \frac{TP + TN}{TP + FN + FP + TN} * 100\% \quad (4.24)$$

D'autres indicateurs évaluent également les performances du classifieur telles que :

b. **La sensibilité**

La sensibilité (TPR), ou True Positif Ratio, qui correspond au taux de vrais positifs. Elle est donnée par :

$$\text{sensibilité} = TPR = \frac{TP}{TP + FN} * 100\% \quad (4.25)$$

c. **La spécificité**

La spécificité (TNR) ou True Négatif Ratio qui correspond au taux de vrais négatifs. Elle est donnée par :

$$\text{spécificité} = TNR = \frac{TN}{TN + FP} * 100\% \quad (4.26)$$

4.4 Tests et résultats de la classification

Les machines à vecteurs supports sont utilisées pour la classification d'ECG réels en deux classes de signaux cardiaques : normaux et anormaux, moyennant les paramètres caractéristiques (caractérisant le comportement des données ECG originales), comme entrées du procédé de classification.

Pour la classification SVM, il est impératif de subdiviser notre base de données en deux groupes d'échantillons, un échantillon pour l'apprentissage pour entraîner le modèle SVM et un autre pour le test afin de le valider. Les approches proposées se définissent par rapport au taux de la précision de cette classification [85]. Afin d'obtenir un taux de précision élevé, nous

utilisons une méthode de la validation croisée (k-fold cross-validation) [115] et nous évaluons statistiquement les performances des méthodes de classification proposées.

La validation croisée est une méthode d'estimation de fiabilité d'un modèle basée sur une technique de séparation de l'ensemble des données en k sous-ensembles où un sous-ensemble sert pour le test et les $k-1$ autres pour l'apprentissage. On itère l'opération sur tous les k sous-ensembles et on calcule ensuite un taux moyen de classification sur la $k^{\text{ième}}$ partie.

Notre système de classification des signaux ECG basé sur la connaissance à priori du nombre de classes, qui est de deux : Normal et Anormal est représenté par le schéma de la figure 4.3.

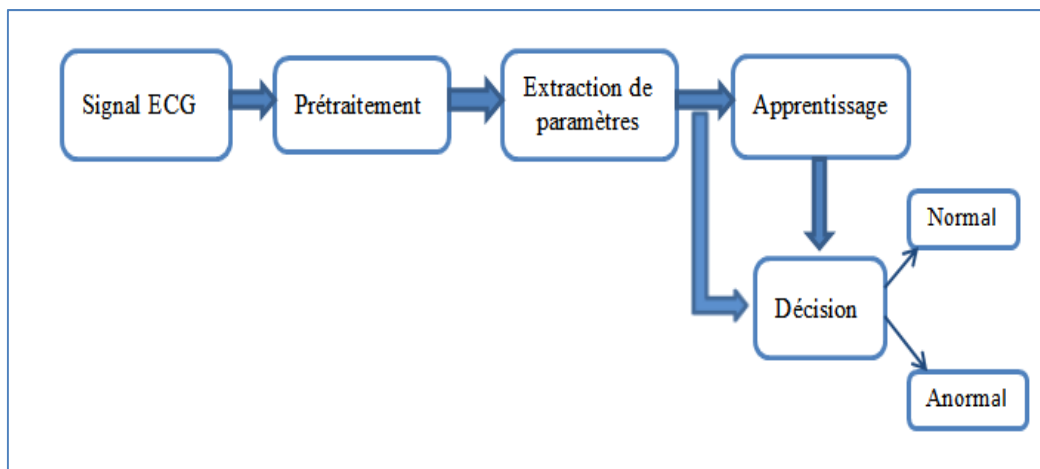


Figure 4.3: Schéma général du système de classification.

Dans nos travaux, la classification des données médicales, se fait à partir des paramètres des gaussiennes sélectionnées des battements des signaux ECG pour la première approche qui est l'approche temporelle et ceux de la variance de l'énergie (fonction statistique) des coefficients de détail des cinq premiers niveaux de décomposition, pour la seconde approche (approche temps-fréquence). Avant d'aborder les différentes méthodes d'extraction de paramètres et de classification, un prétraitement numérique des signaux ECG de longue durée (Holter) a consisté à éliminer la ligne de base à l'aide de filtres numériques et de définir chaque battement d'un signal par le processus de découpage, (voir chapitre 2 paragraphe 2.3.1).

4.4.1 Résultats de classification moyennant RBF, GOFR et SVM

En exploitant la base de données universelle MIT-BIH, deux bases de données ont été établies: l'une pour l'apprentissage, l'autre pour le test du classificateur proposé. Les signaux ECG choisis sont issus de la MIT-BIH arrhythmia data base où trois pathologies différentes

(Arythmie, arythmie supraventriculaire, extrasystole ventriculaire) sont utilisées ainsi que des signaux ECG normaux.

L'apprentissage est l'étape la plus importante dans la réalisation d'un système de détection automatique des anomalies cardiaques. Quand elle se déroule dans de bonnes conditions, cela permet d'avoir de bonnes performances du système de classification. C'est pour cette raison que l'élaboration de la base de données est très importante. Notre travail comprend en deux étapes :

1. L'extraction des paramètres caractérisant le signal.
2. Classification des signaux «Battements» à l'aide du classificateur **SVM**.

4.4.1.1 Extraction de paramètres temporels moyennant RBF et GOF

Notre banque d'apprentissage se compose de 364 battements normaux et 1148 battements de cœur anormaux (Arythmies ventriculaires, arythmies supra-ventriculaires, extrasystoles ventriculaires). Pour développer l'apprentissage SVM, 1208 battements cardiaques ont été prélevés au hasard à partir de 1512 battements d'apprentissage et 300 battements pour tester notre classifieur proposé.

Chaque battement est modélisé par l'algorithme GOF (Generalized Orthogonal Forward Regression). L'objectif de la phase de modélisation par le GOF dans cette application est l'extraction de caractéristiques A_i^* , σ_i^* , μ_i^* obtenus après sélection et ajustement pour chaque fonction Gaussienne (chapitre2). Le modèle de chaque battement est une combinaison linéaire des cinq fonctions Gaussiennes qui représentent chaque onde caractéristique du battement cardiaque [1].

4.4.1.2 Classification des signaux ECG par SVM:

Dans nos travaux, on s'est limité au problème de discrimination à deux classes « discrimination binaire » c'est-à-dire $y \in \{-1, 1\}$.

Il est décidé que x est de classe :

- 1 si $f(x) \geq 0$.
- -1 dans le cas contraire.

La frontière de décision $f(x) = 0$ est un hyperplan appelé hyperplan séparateur. Dans notre cas, il n'existe pas de séparateur linéaire. On reconsidère le problème dans un espace de dimension supérieure. On utilise une astuce qui consiste à utiliser une fonction noyau K dont le but est de faire le calcul dans l'espace d'origine, ou $K = \Psi(x_i)^T \Psi(x_j)$.

En pratique la fonction doit respecter certaines conditions :

- Elle doit correspondre a un produit scalaire dans un espace de grand dimension.
- Elle doit être symétrique.

Il existe plusieurs noyaux pour l'usage des SVM telles que :

- Fonction gaussienne ou RBF « **Radial Basis Functions** ».
- Fonction polynomiale : $K(x_i, y_j) = (x_i^T y_j)^d$.
- Fonction noyau linéaire : $K(x_i, y_j) = x_i^T y_j$.

L'organigramme qui décrit le principe de l'apprentissage est représenté dans la figure 4.4 suivante :

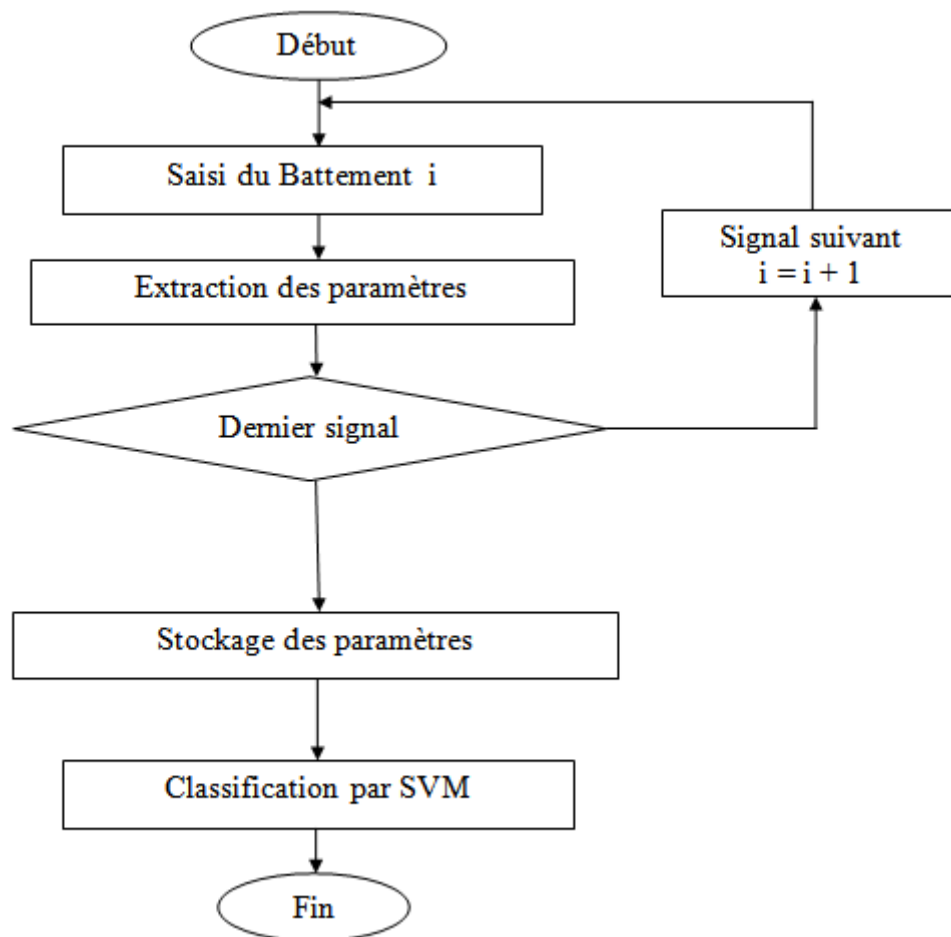


Figure4.4 : Principe de l'apprentissage.

L'étape de la reconnaissance est plus simple que celle de l'apprentissage, il suffit d'introduire les caractéristiques du signal à reconnaître qui permettent de calculer la sortie y qui correspond au

vecteur d'entrée x à l'aide de la fonction discriminante linéaire $f(x)$ et le résultat obtenu par la validation de la classe associé au signal :

- 1 si signal Normal.
- -1 si signal Anormal.

Le processus de reconnaissance est illustré par le schéma de l'organigramme de la figure 4.5.

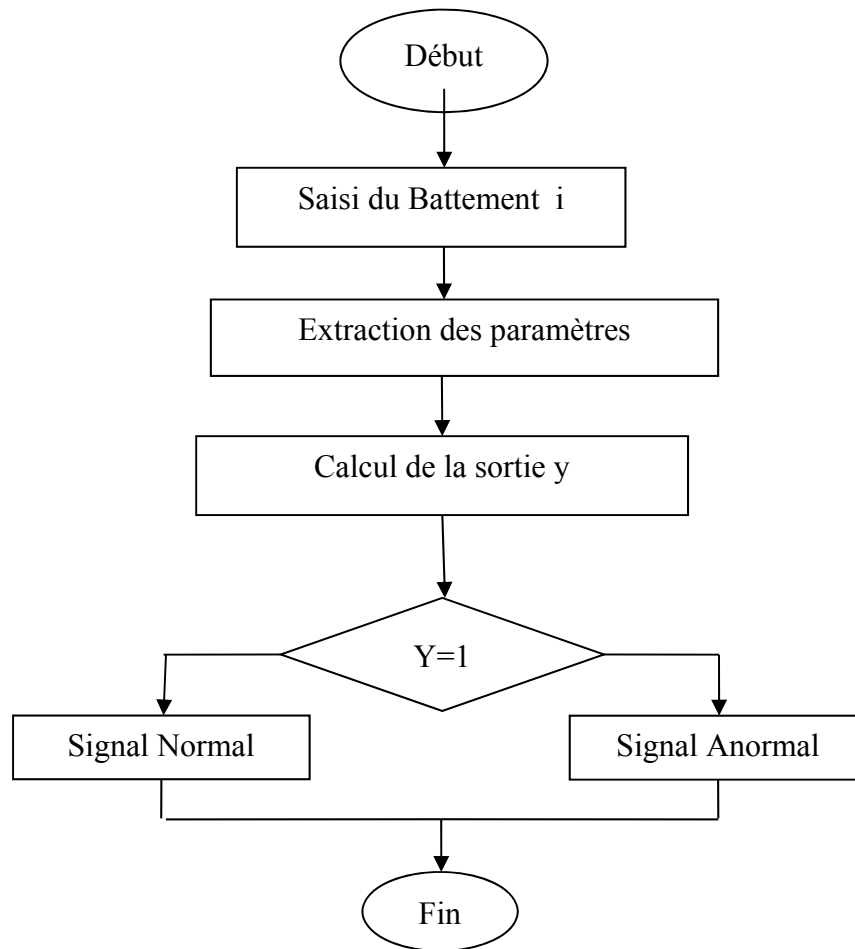


Figure 4.5 : Organigramme de la reconnaissance.

La classe de distribution des échantillons pour l'apprentissage et le test sont résumés dans le tableau 4.2.

Tableau 4.2 : Classes de distribution des échantillons pour l'apprentissage et le test

Classe	Apprentissage	Test	Total
Normal	214	150	364
Anormal	998	150	1148

En outre, il est évident que l'utilisation des paramètres du noyau approprié est cruciale pour obtenir de bonnes performances de classification. Pour une bonne performance de l'apprentissage SVM, nous choisissons les fonctions de base radiales (RBF) et essayons de trouver les paramètres du noyau appropriés σ , et γ . Les valeurs optimales de σ , et γ ne peuvent être déterminées qu'après essais de différentes valeurs et obtention du meilleur résultat.

Les résultats [1] de la classification des battements cardiaques moyennant les algorithmes GOFR pour la sélection des paramètres et SVM pour la classification proprement dite, sont illustrés ci-dessous. Pour une base de données de 364 battements normaux et 1148 anormaux, nous obtenons les matrices de confusion de l'ensemble d'apprentissage (Tableau 4.3) et de l'ensemble Test (Tableau 4.4).

Tableau 4.3 : Matrice de confusion de l'ensemble d'apprentissage

	Battement Normal	Battement Anormal
Battement Normal	214	0
Battement Anormal	0	998

Les résultats du tableau 4.3, prouvent l'efficacité du classificateur proposé à l'étape d'apprentissage où la sensibilité, la spécificité, et la précision de notre classificateur sont de 100%. L'étape de test est résumée dans le Tableau 4.4.

Tableau 4.4: Matrice de de confusion de l'ensemble de Test

	Battement Normal	Battement Anormal
Battement Normal	148	2
Battement Anormal	3	147

Afin d'évaluer les résultats de la phase d'entraînement, la performance du classifieur (Tableau 4.5) est déterminée par les paramètres statistiques qui sont calculés par la matrice de confusion.

Tableau 4.5 : Performance du classifieur proposé.

Performance	Apprentissage	Test	Validation
TPR %	100	98,01	99,17
TNR %	100	98,66	99,82
Tc %	100	98,33	99,67

Les résultats du tableau 4.4 reflètent la performance du classifieur proposé qui est de 99.67%.

Pour une meilleure fiabilité de notre système de classification, nous appliquons la méthode de 10-fold cross validation. L'ensemble des battements est divisé en 10, ensuite on sélectionne un des dix échantillons comme un ensemble pour la validation et les 90% autres échantillons constitueront l'ensemble d'apprentissage. On s'assure que chaque sous-échantillon ait été utilisé exactement une fois comme ensemble de validation. Nous obtenons les résultats du tableau 4.6 qui donnent les pourcentages de la classification basée sur la combinaison RBF- GOFR et utilisant trois paramètres caractéristiques (l'écart type, l'amplitude et la moyenne des gaussiennes). La précision moyenne sur les dix estimations obtenues à partir des résultats des tests est considérée comme le critère d'évaluation des performances de l'algorithme de classification.

Tableau4.6 : Performances de l'approche temporelle

10-Fold	TPR%	TNR%	TC%
Test1	98,01	98,66	98,83
Test2	99,17	99,82	99,67
Test3	100	100	100
Test4	100	99,56	99,87
Test5	98,78	99,93	99,84
Test6	99,01	100	99,85
Test7	99	100	99,75
Test8	99	98,95	98,93
Test9	99,64	99,76	99,88
Test10	98,39	99,86	99,56

La figure 4.6 représente la distribution de TPR, TNR et TC à chaque itération de la validation croisée. Cette méthode a permis d'obtenir des valeurs optimales de 99,17%, 99.82% et 99,67% concernant, respectivement, la sensibilité, la spécificité, et la précision de classification de nos données d'enregistrement ECG .

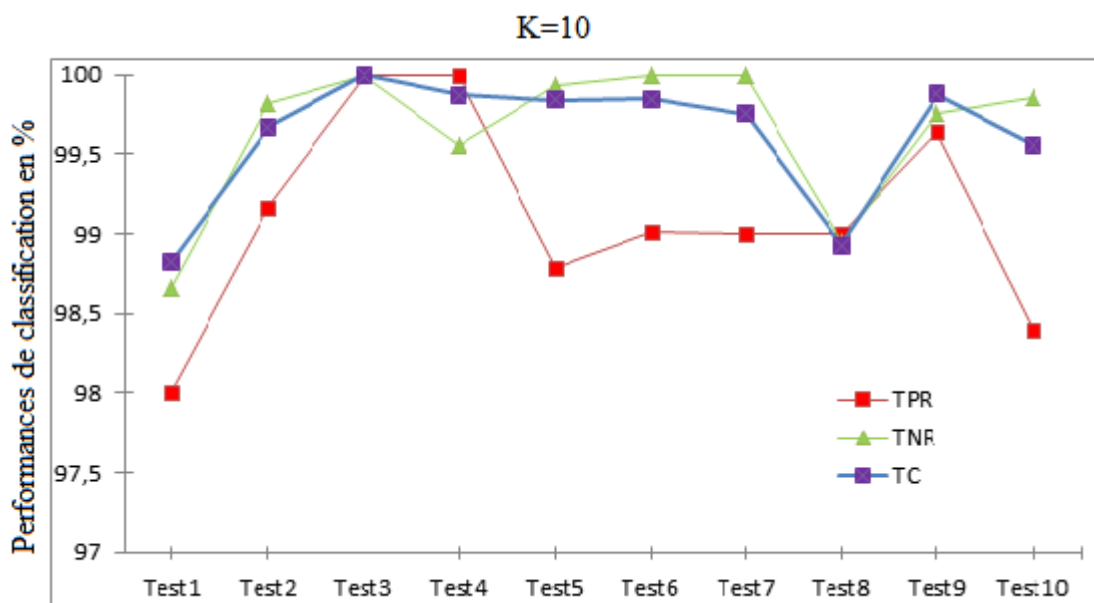


Figure 4.6 : Performances de classification de l'approche RBF- GOFr-SVM avec la validation croisée

Nos résultats démontrent que la performance des tests du réseau de neurones et du système de classification SVM, est très satisfaisante puisque un taux de reconnaissance de 99.67% est obtenu.

Il serait intéressant de comparer notre méthode avec d'autres systèmes de reconnaissance présentés dans la littérature. Avant cela, nous proposons une comparaison avec une seconde méthode qui est l'approche Temps-Fréquence.

4.4.2 Résultats de classification moyennant DWT, ACP, ACI et SVM

4.4.2.1 Extraction de paramètres statistiques moyennant la DWT

Comme pour la première approche, les signaux ECG anormaux choisis pour l'apprentissage et le test, sont issus de la MIT-BIH arrhythmia database où trois pathologies différentes (Arythmie, arythmie supraventriculaire, extrasystole ventriculaire) sont prises ainsi que des signaux ECG normaux. Nous avons donc exploité un ensemble d'enregistrements incluant deux bases de signaux ECG. L'une consiste en $N1 = 160$ ECG normaux et l'autre en $N2 = 160$ ECG pathologiques. Leur durée est de 10 s et la fréquence d'échantillonnage est de 128 Hz.

Pour tous les signaux normaux et pathologiques, nous appliquons les DWT les plus utilisées dans le traitement des signaux biomédicaux ECG telles que les ondelettes de Coiflet, Daubechies ... etc. Nous déterminons les différents coefficients de détail pour les cinq(5) premiers niveaux de décomposition puis nous calculons l'énergie des coefficients de détail correspondant à chaque niveau de décomposition. Nous calculons par la suite, les variances d'énergie des cinq premiers niveaux de décomposition. Comme la fréquence d'échantillonnage est de 128Hz, les bandes de fréquences des coefficients de détail seront D1 (32-64 Hz), D2 (16-32 Hz), D3 (8-16 Hz), D4(4-8 Hz) et D5 (2-4 Hz) [1]. Les bandes de fréquence correspondant aux différentes ondes du signal de l'ECG sont: QRS (5-15 Hz), T (0.1-4 Hz) et P (1-5 Hz). Donc, la sous-bande de fréquence du premier niveau de détail D1 est en dehors des caractéristiques des ondes de l'ECG. Le troisième niveau D3 contient les informations du complexe QRS. Les quatrième et cinquième niveaux D4 et D5, incluent des informations sur les ondes T et P.

Les coefficients de détail sont donc utilisés pour caractériser l'ECG dans le domaine temps - fréquence et véhiculer les informations des différentes ondes qui le caractérisent.

Notre algorithme est appliqué sur 320 signaux ECG réels normaux et anormaux parmi lesquels nous prendrons deux seulement, afin d'illustrer quelques résultats :

- Le premier signal ECG de la figure 4.6 est un signal normal noté **ECGn**, avec des caractéristiques bien définies et un cycle cardiaque régulier (16420).
- Le deuxième signal ECG de la figure 4.7 est un signal présentant une arythmie supra ventriculaire (ESV S84-812m) et est noté **ECGan**.
-

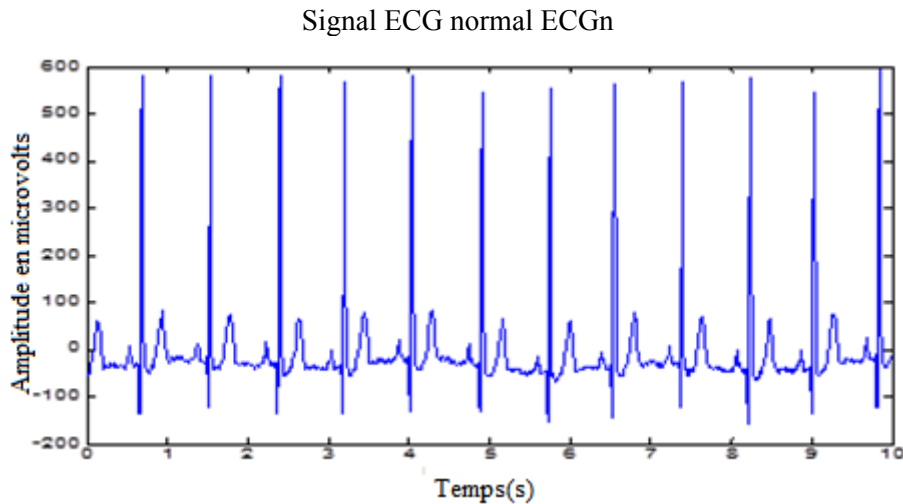


Figure4.7 : Signal ECGn d'un sujet sain (16420).

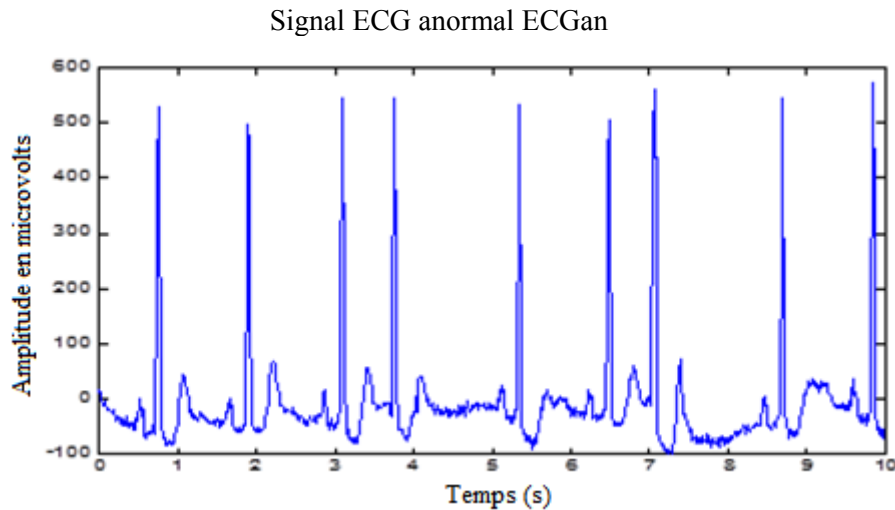


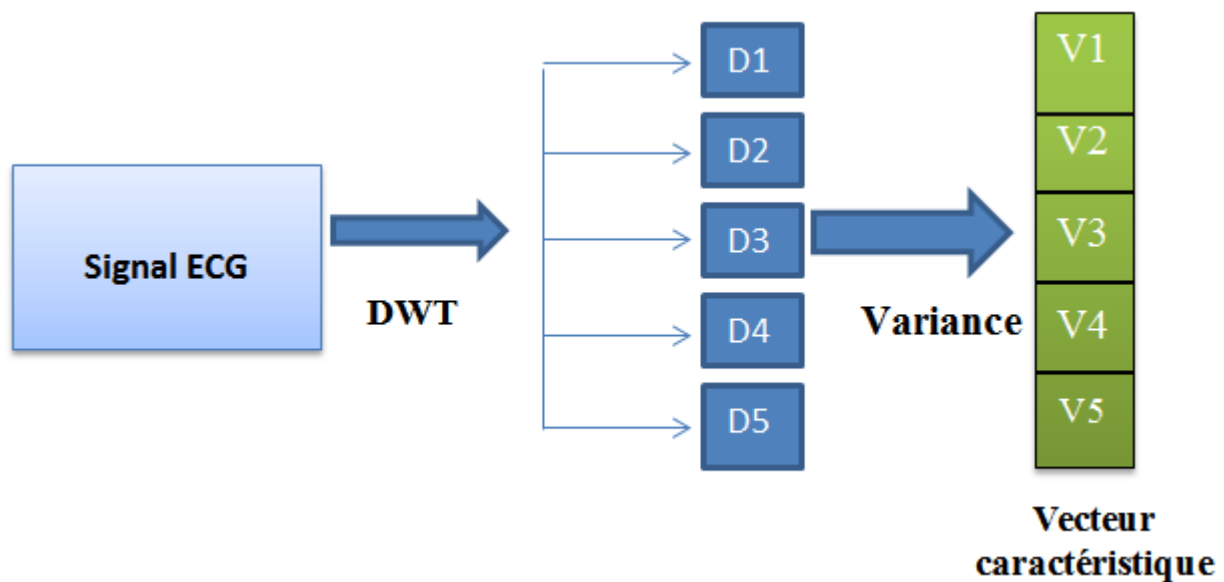
Figure4.8 : Signal ECGan d'un sujet pathologique (ESV812m).

Les signaux ECG issus de la banque de données, qui serviront à l'apprentissage, sont tous échantillonnés à la même fréquence d'échantillonnage $F_e=128\text{hz}$. Ainsi les filtres FIR, utilisés par la DWT seront dans les bandes de fréquences données dans le tableau4.7.

Tableau 4.7: Bandes de fréquences des différents niveaux de décomposition

Coefficients de détails	Bandes de fréquence(Hz)
D1	32–64
D2	16–32
D3	8–16
D4	4–8
D5	2–4

L'application de la décomposition par DWT en utilisant **l'ondelette de coifly2**, aux signaux ECG, permet la décomposition de chaque signal en cinq niveaux (N= 5) de détail, (figure 4.9).

**Figure 4.9 :** Extraction des paramètres caractéristiques d'un signal ECG.

- **Représentation des coefficients de détail (CDi):**

L'application de la DWT de Coif2, sur les signaux ECGn et ECGan, nous donne les coefficients d'approximation et de détail représentés sur les figures 4.10 et 4.11.

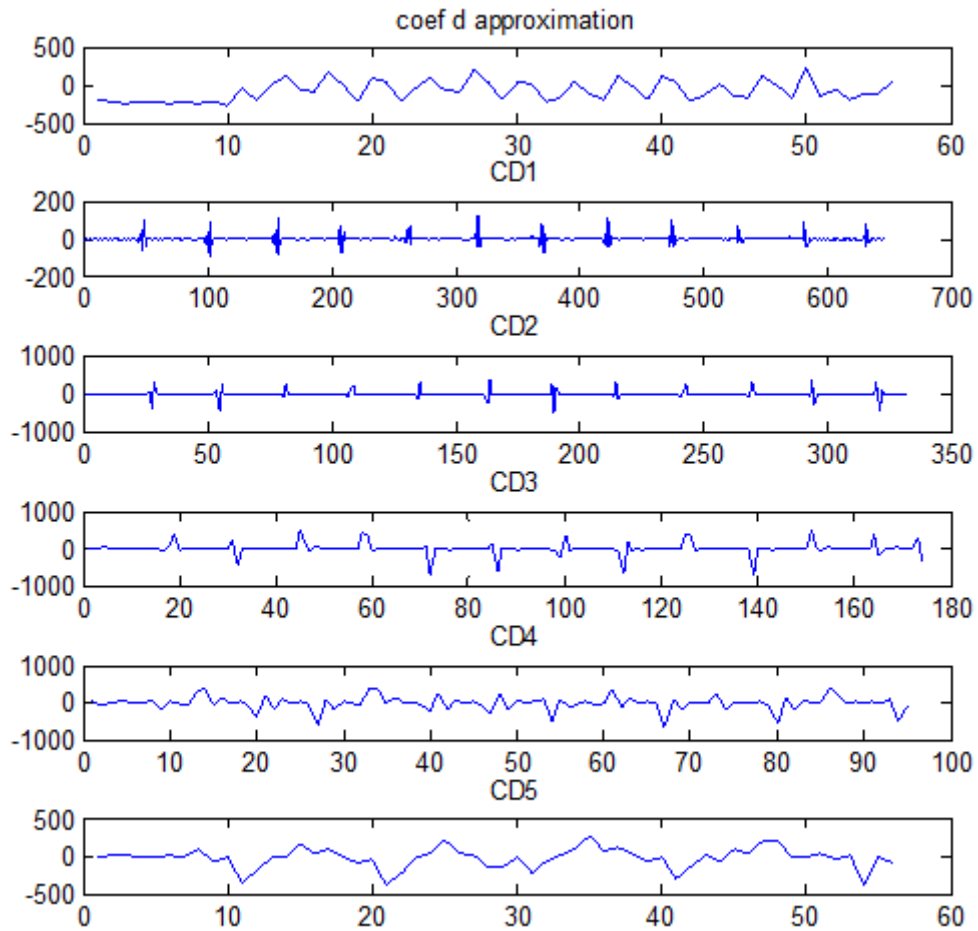


Figure 4.10 : Coefficients d'approximation et de détail pour le signal normal ECGn 16420(5 niveaux).

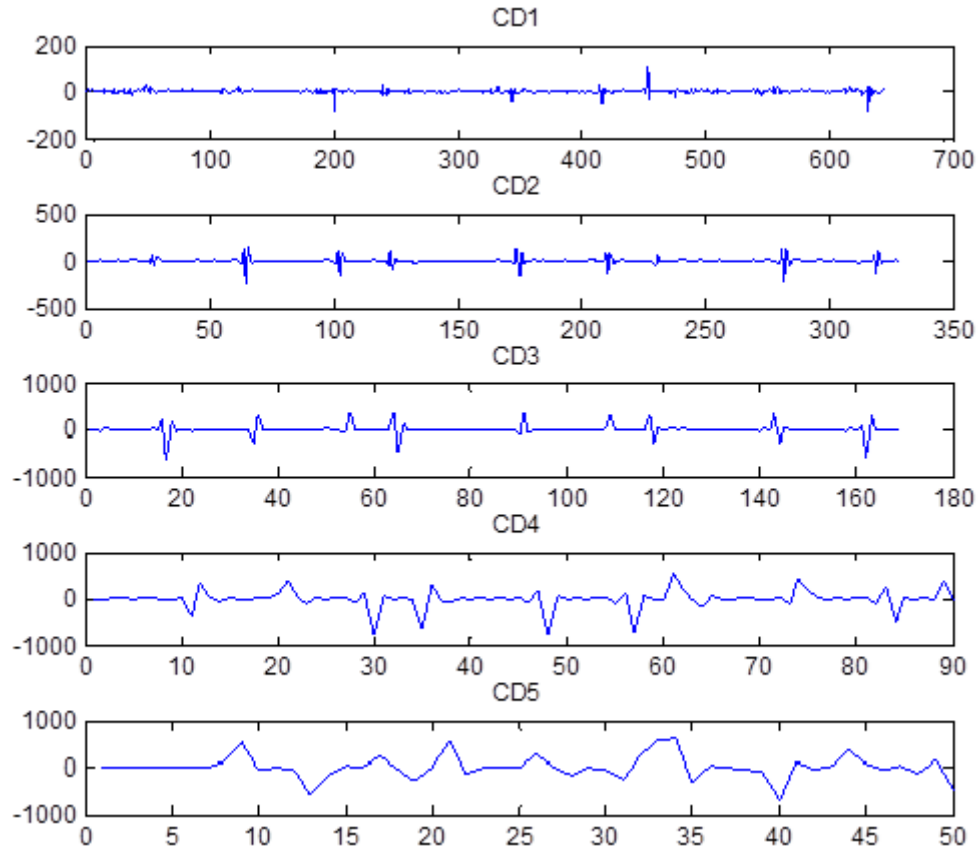


Figure4.11 : Coefficients de détail pour un signal anormal ECGan (5 niveaux).

- **Calcul de la variance d'énergie :**

Après l'application de la DWT sur les signaux ECG, nous allons calculer la variance d'énergie des coefficients de détails. Pour chaque signal de coefficient de détail CD_j , de taille N de moyenne $\overline{CD_j}$, on définit la variance d'énergie Var:

$$\text{Var} = \frac{1}{N} \sum_{i=1}^N \|(CD_j)_i - \overline{CD_j}\|^2 \quad (4.27)$$

j : Nombre de niveau de décomposition.

N : représente le nombre des échantillons des coefficients de détail (CD_j).

Ou d'une manière plus générale elle s'écrit :

$$\text{Var} = \frac{1}{N} \sum_{i=1}^N \|x_i - \bar{x}\|^2$$

Cette mesure permettra de quantifier les différentes variations énergétiques se produisant au niveau des bandes de fréquences de signal ECG et montrer ainsi que l'information énergétique du signal ECG est traduite par les différents coefficients de détails.

Afin de voir l'impact de la distribution des énergies des signaux normaux et pathologiques sur les différents niveaux de décomposition, nous avons établi les histogrammes en figures 4.12 et 4.13 donnant le pourcentage d'énergie en fonction des différents niveaux.

Nous constatons que pour le signal normal, le pourcentage d'énergie du coefficient de détail dans le 3^{ème} niveau est beaucoup plus important (plus de 30% de l'énergie totale) que dans les autres niveaux (figure 4.12). En revanche, pour le signal pathologique, le pourcentage de l'énergie est maximal au niveau 4 ou 5 (figure 4.13).

L'information énergétique du signal est concentrée dans les coefficients de détails à plus de 90%. Le niveau 3 de la bande de fréquence (8-16Hz) correspondant au QRS paraît pertinent pour les signaux normaux. Le cas pathologique montre une augmentation des niveaux 4 et 5 aux dépens des niveaux 2 et 3. Donc perturbation au niveau de la basse fréquence du signal ECG: l'une des ondes T ou P est donc touchée par l'anomalie.

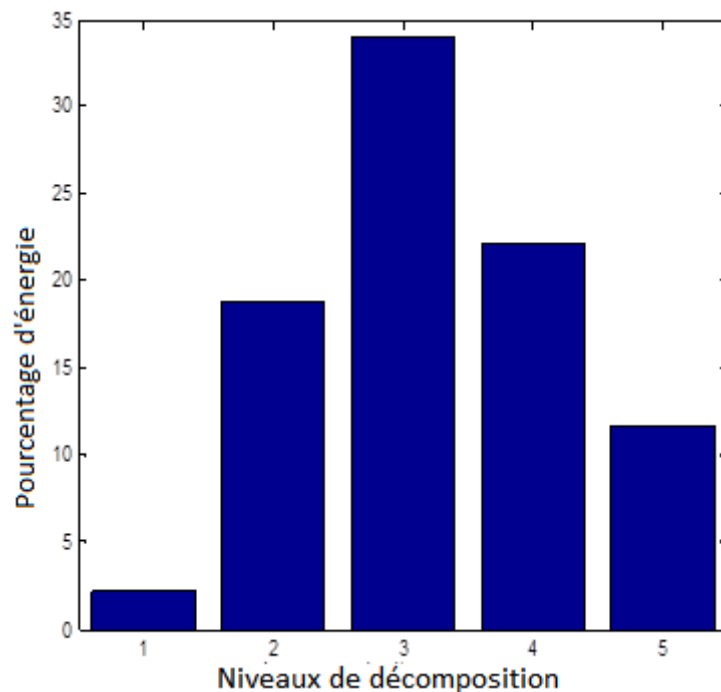


Figure 4.12 Pourcentage d'énergie du signal normal ECGn en fonction de chaque niveau de décomposition.

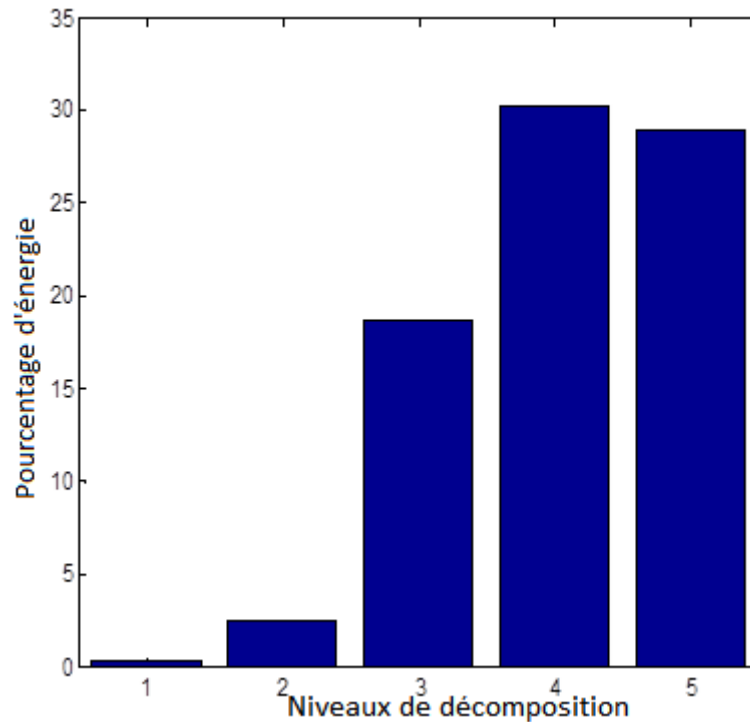


Figure 4.13: Pourcentage d'énergie du signal anormal ECGan en fonction de chaque niveau de décomposition.

Pour avoir une idée à priori sur la distribution des énergies de tous les signaux normaux et pathologiques sur les différents niveaux de décomposition, nous avons établi des histogrammes en figures 4.14 et 4.15 donnant le pourcentage d'énergie pour 11 signaux ECG quelconques en fonction des différents niveaux.

Notons que pour chaque sujet, une colonne de l'histogramme représente, de gauche vers la droite, les énergies de variances de D1 à D5 (i.e., des plus hautes fréquences vers les plus basses). La plus grande partie de l'énergie (90%) est concentrée dans les cinq premiers niveaux des coefficients de détail, c'est pourquoi nous avons limité notre décomposition à 5 niveaux. Pour tous les sujets sains (Figure 4.14), nous remarquons une prédominance de l'énergie de variance du niveau D3 de bande de fréquence (8-16Hz). Les cas pathologiques sont caractérisés par une prédominance des niveaux 4 et 5 (Figure 4.15).

A cette étape de notre travail, nous pouvons conclure que la variance d'énergie est un paramètre pertinent pour caractériser le cas normal du cas anormal. Les variances d'énergie des cinq premiers niveaux de décomposition de la DWT représentent les paramètres pertinents des signaux ECG. Les coefficients de détail présentent le maximum d'énergie du signal (90% d'où leur pertinence).

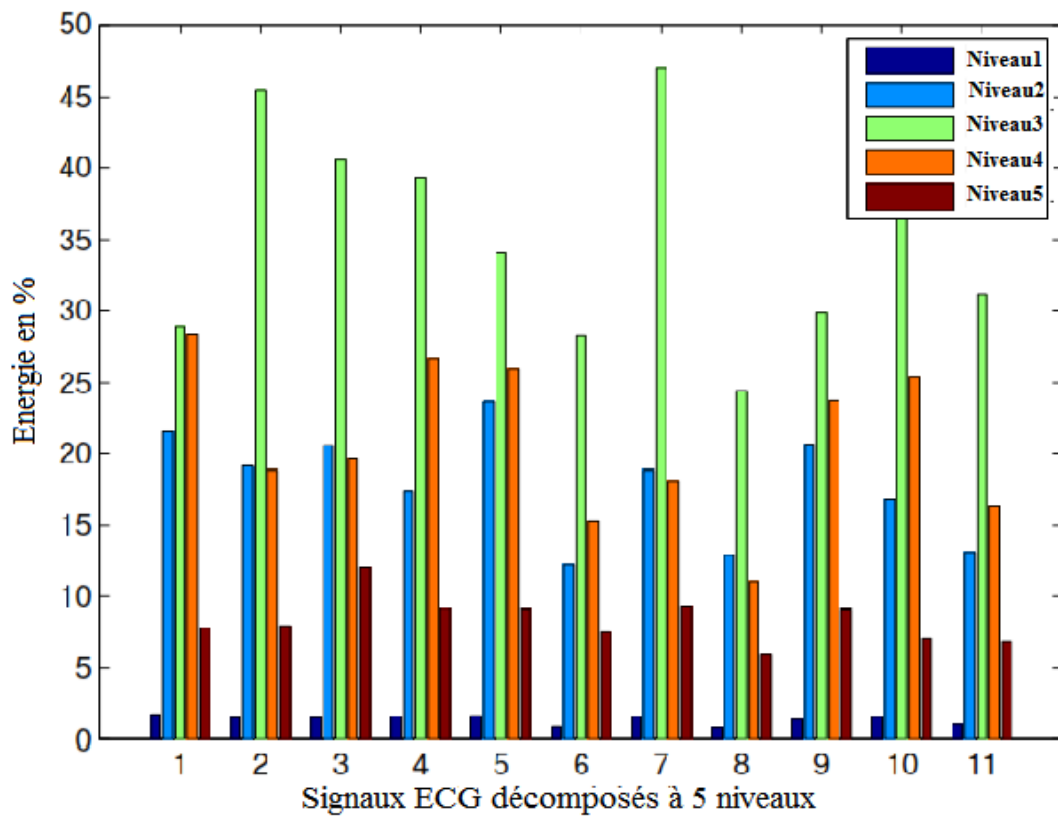


Figure 4.14 : Histogramme de la variance d'énergie des coefficients de détail avec la DWT de Coiflet pour 11 sujets sains.

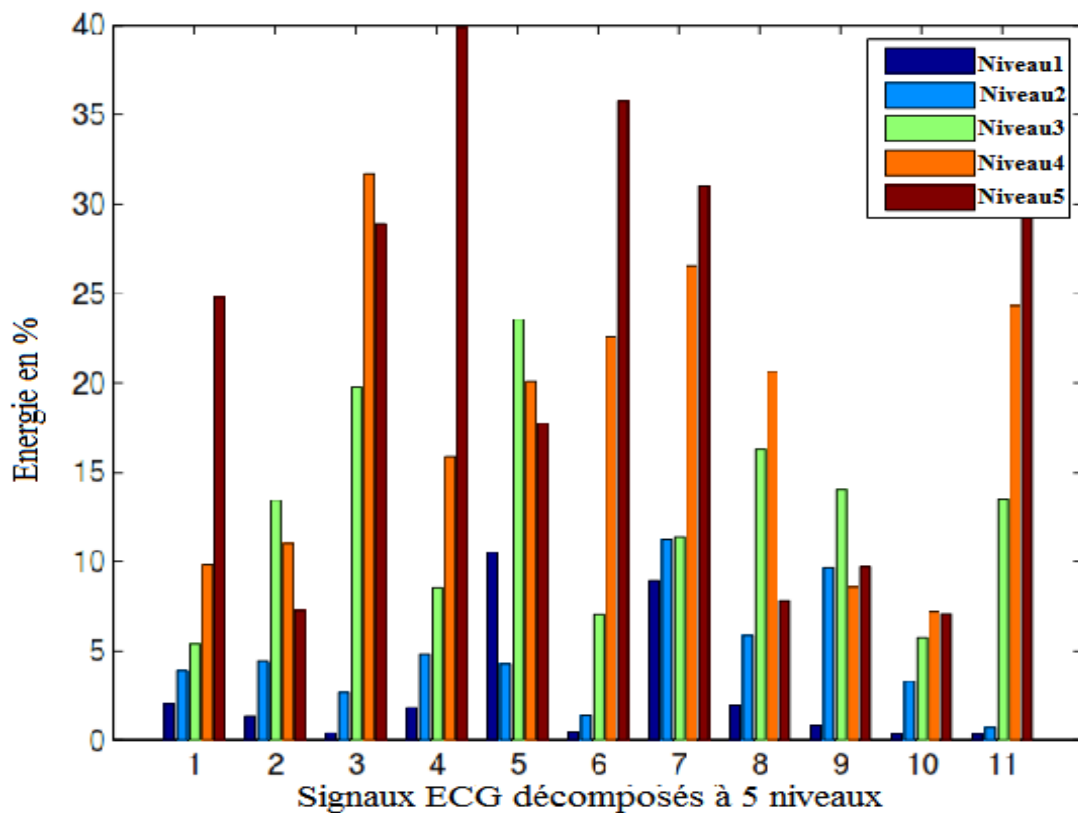


Figure 4.15: Histogramme de la variance d'énergie des coefficients de Détail avec la DWT de Coiflet pour 11 signaux ECG anormaux.

Une étude statistique faite sur l'ensemble des données de la banque montre les mêmes caractéristiques. Notre descripteur du signal ECG sera donc la variance énergétique des coefficients de détail des 5 premiers niveaux de décomposition de la DWT

4.4.2.2 Réduction de dimension

Les matrices d'ACP et d'ACI procèdent à la réduction de dimensions de l'espace de mesure en gardant les paramètres les plus pertinents sans les classer. Le vecteur caractéristique étant à 5 dimensions, il est difficile de faire sa représentation. Afin de réduire cet espace, on a recours à la méthode de l'analyse en composante principale ou indépendante. L'application de l'ACP ou ACI, n'est autre qu'un changement d'espace où l'on peut choisir les axes de projection, pour garder les plus pertinents.

Après l'application de la DWT, nous calculons la variance de l'énergie des coefficients de détail d'ondelettes de chaque niveau de décomposition et pour chaque signal ECG normal ou pathologique. Nous obtenons deux matrices, l'une pour les signaux normaux $Nor(N_1, p_1)$ et l'autre pour les signaux anormaux $Anor(N_2, p_2)$, tels que :

N_1 : le nombre de lignes représentant le nombre de signaux normaux (160).

p_1 : le nombre de colonnes représentant le nombre de niveaux(5), (une colonne pour chaque niveau).

N_2 : le nombre de lignes représentant le nombre des signaux anormaux (160)

p_2 : le nombre de colonnes représentant le nombre de niveaux(5), (une colonne pour chaque niveau).

Nous nous retrouvons avec une base de données $(N_1 \times 5 + N_2 \times 5)$ que l'on doit réduire par ACP ou ACI.

Les valeurs d'énergie calculées de chaque coefficient, sont mises dans une matrice d'énergie $E(i, j) = [Nor, Anor]$ qui englobe les énergies des coefficients de détails des cinq niveaux de décomposition des signaux normaux et anormaux. Cette matrice subit des transformations matricielles telles que : la diagonalisation qui permet de déterminer les axes de projection qui ont la plus grande part d'inertie., extraction des valeurs propres (λ_i) et matrice des vecteurs propres (u), qui vont caractériser les nouveaux axes. Ce qui constituera la matrice d'ACP V_{Acp} proprement dite et donc permettra de déterminer les principaux axes de l'ACP, afin de décorréliser les signaux normaux des pathologiques. A partir de la distribution des valeurs propres (variance) de la matrice contenant les signaux à classer, (figure 4.16) on peut remarquer que la majorité de la variance des vecteurs caractéristiques de ces signaux est concentrée sur les deux premiers axes.

Nous avons donc choisi les deux derniers axes car la dimension de projection $P= 96\%$ dépasse 80%.

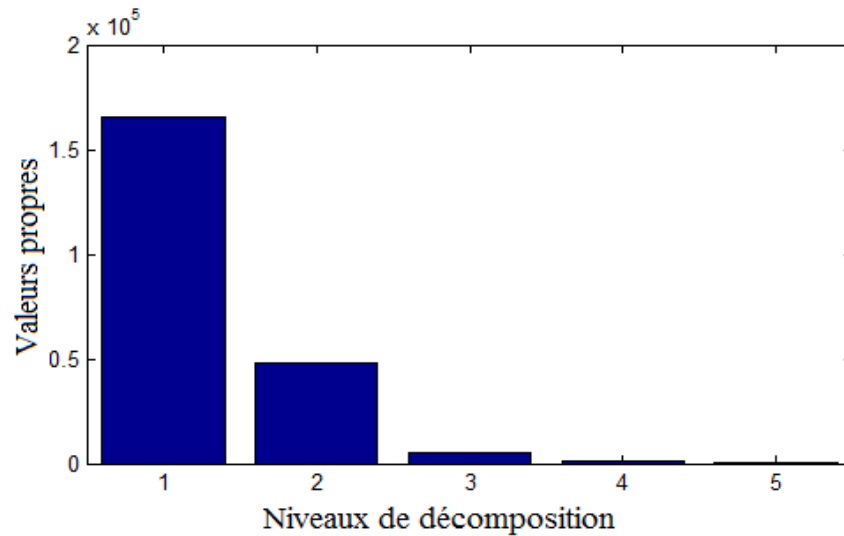


Figure 4.16 : Représentation des valeurs propres de la base de données.

Nous proposons donc de réduire la dimension des vecteurs caractéristiques à deux dimensions, afin de minimiser les calculs, mais aussi pour pouvoir bien visualiser l'espace de projection. Le résultat de l'ACP est représenté dans la figure 4.17.

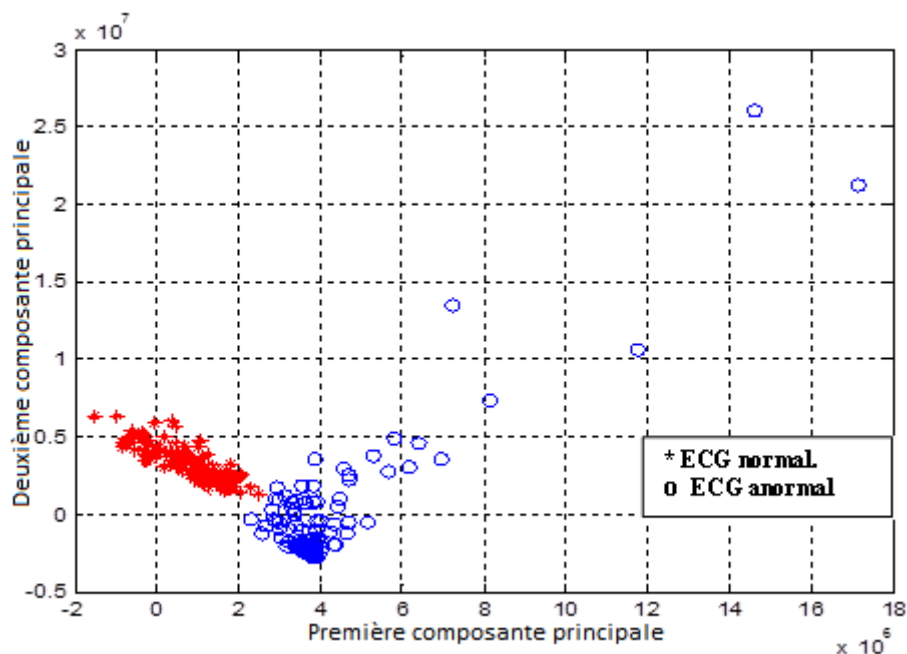


Figure 4.17 : Nuage des points normaux et anormaux représentés sur les deux composantes principales de l'ACP.

Nous remarquons dans cette dernière figure que le nuage des signaux normaux et celui des signaux anormaux sont bien distincts. C'est l'objectif de l'application en composante principale ACP.

De la même manière que pour l'ACP, la projection de nos données sur les principaux axes de l'ACI permet d'observer deux régions distinctes représentant les deux classes à séparer, figure (4.18). L'appartenance de ces deux régions de points est connue à priori et servira comme données d'apprentissage dans la méthode de classification.

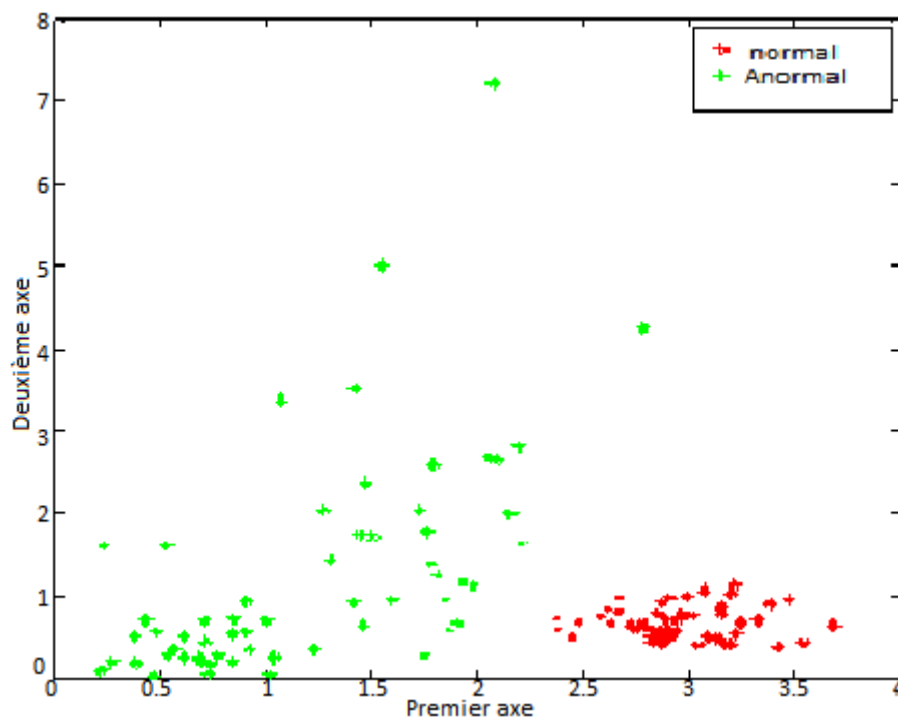


Figure 4.18: Projection sur les deux axes de l'ACI (DWT Rbio3.9).

Après application de L'ICA nous obtenons les projections des paramètres sur les deux axes principaux de l'ICA. Nous observons deux régions distinctes représentant les deux classes à séparer. Pour définir les limites des deux régions de classes déterminées dans la phase précédente, nous appliquons l'algorithme SVM qui permet de trouver l'hyperplan de séparation.

4.4.2.3 Apprentissage et Classification :

L'application de l'ACI, nous a permis de distinguer les deux différentes classes, la méthode de classification par SVM, nous permettra de définir les limites des deux régions de classes, et sera nécessaire pour la prise de décision afin de classer un signal ECG de classe inconnue.

Les résultats [1] de la classification des battements cardiaques moyennant les algorithmes **DWT**, **ACP**, **ACI** et **SVM** seront illustrés ci-dessous. La distribution des signaux ECG est illustrée dans le tableau 4.8.

Tableau 4.8 : Distribution des échantillons pour l'apprentissage et le test

Signaux	Classe	Apprentissage	Test	Total
ECG	Normaux	80	80	160
	Anormaux	80	80	160

Afin d'évaluer les résultats de la phase d'apprentissage, la performance du classifieur est déterminée par les paramètres statistiques Tc, TPR, TNR calculés par la matrice de confusion comme pour la première approche. La performance de la SVM a été réalisée en variant le noyau jusqu'à l'obtention du meilleur taux de reconnaissance.

✚ En appliquant la SVM pour un noyau linéaire, il apparaît clairement l'inexistence d'une séparation linéaire optimale, (figure4.19).

Les propriétés importantes des SVM nous permettent de faire une séparation non linéaire des classes.

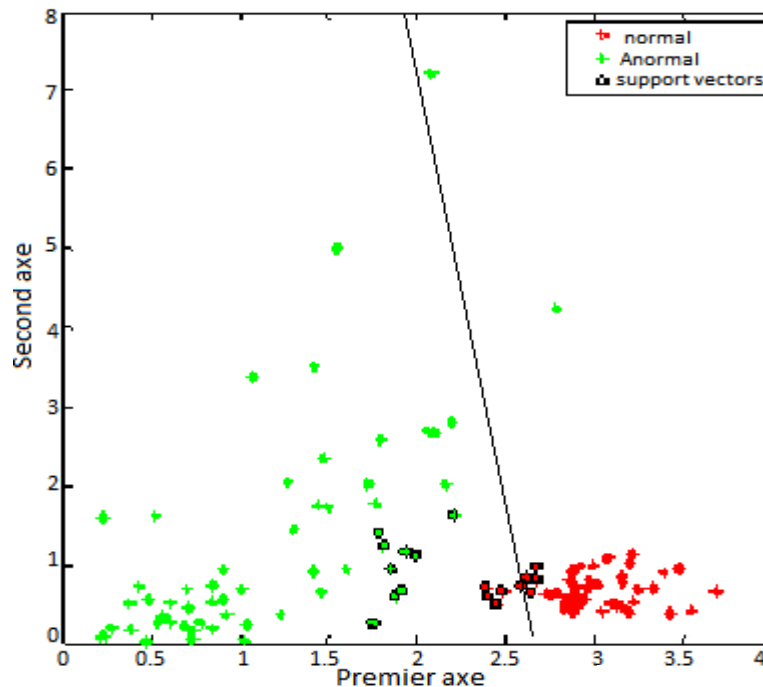


Figure4.19: Cas de discrimination linéaire.

La fonction noyau utilisée étant une fonction de type polynomial d'ordre deux. Nous obtenons le résultat qui apparaît sur la figure 4.20.

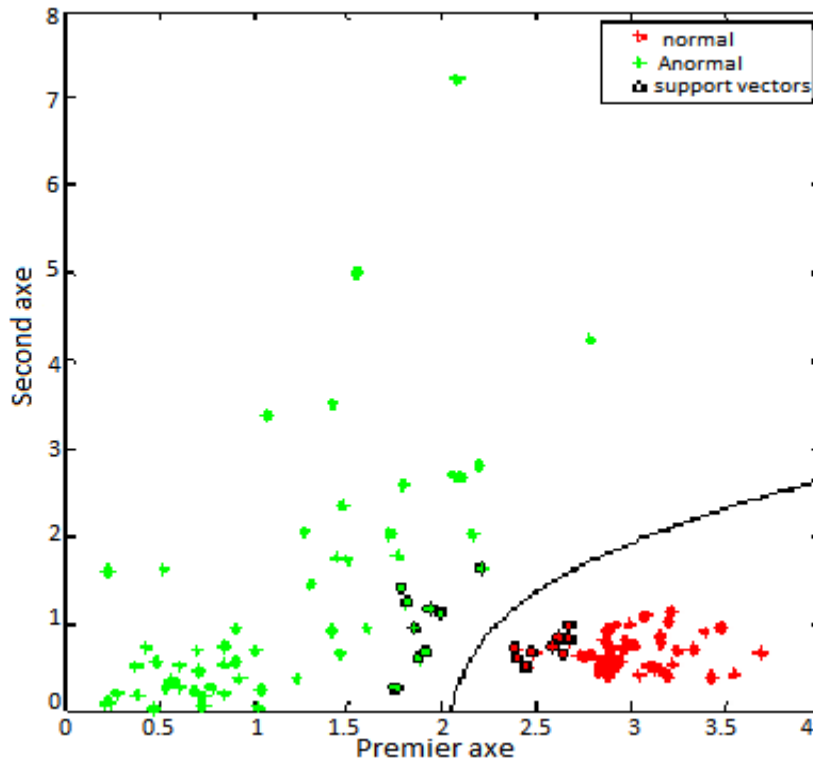


Figure 4.20 : Classification par SVM avec Rbio3.1 et noyau polynomial du 2nd ordre.

Les résultats à l'étape d'apprentissage sont très satisfaisants puisque les paramètres de performance, à savoir la sensibilité, la spécificité, et la précision de notre classifieur sont de 100%. A l'étape de test, la performance du classifieur est déterminée par les paramètres statistiques calculés par la matrice de confusion. Un taux de classification de 100% a été obtenu. Il faut noter que pour le même noyau nous avons testé l'ensemble des DWT. Nous avons étudié l'impact du choix de type d'ondelette sur le taux de classification et réalisé une étude comparative entre les différentes ondelettes issues de la DWT, dans le but de sélectionner celles les plus appropriées pour notre application qui est la classification. Le résultat de comparaison est donné dans le tableau 4.9.

Tableau 4.9 : Performances de classification par SVM de quelques ondelettes

Ondelette	Db1	Db2	Db3	Db4	Db5	Db6	Db7	Db8
Taux classification %	99.93	100	99.16	100	99.58	99.58	99.16	100
Ondelette	Db9	Db10	Coif1	Coif2	Coif3	Coif4	Coif5	Sym1
Taux Classification %	98.75	99.58	98.07	100	98.75	100	98.33	99.16
Ondelette	Sym2	Sym3	Sym4	Sym5	Sym6	Sym7	Sym8	Bior1.3
Taux classification%	99.58	99.16	98.75	99.75	100	99.58	99.58	100
Ondelette	Bior1.5	Bior2.2	Bior2.4	Bior2.6	Bior3.1	Bior3.3	Bior3.5	Bior3.7
Taux classification%	99.5	96.6	98.7	99.16	100	100	97.5	97.5
Ondelette	Bior3.9	Bior4.4	Bior5.5	Bior6.8	Rbio1.3	Rbio1.5	Rbio2.2	Rbio2.4
Taux classification%	92.5	99.5	100	99.5	99	100	97.9	99

D'après le tableau 4.9, nous pouvons constater l'importance du choix d'ondelette sur le taux de classification. Les résultats obtenus ont été satisfaisants pour toutes les ondelettes utilisées. Les ondelettes Db2, Db8, Coif2, Coif4, Bior3.1 et Bio3.3, Bior5.5, Rbior1.5, Sym6 et le classificateur SVM ont permis d'atteindre le plus haut score de classification c'est-à-dire 100%. Les figures 4.21, 4.22, 4.23 et 4.24 illustrent quelques résultats obtenus précédemment où une nette séparation par SVM est obtenue avec un score de 100%.

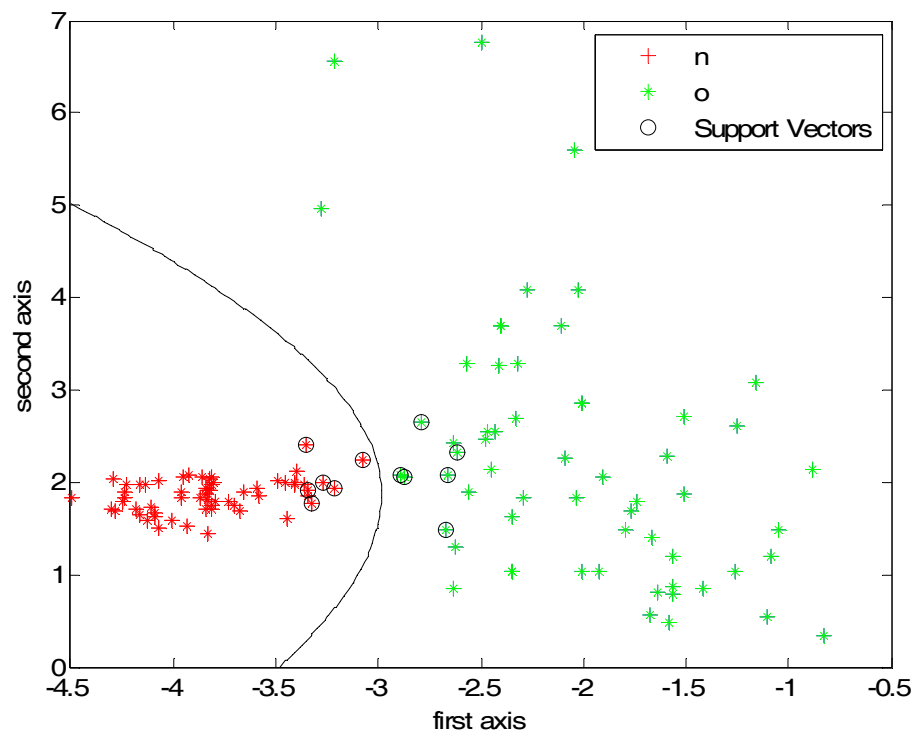


Figure 4.21 : Classification par SVM avec Coif2 et noyau polynomial du 2nd ordre.

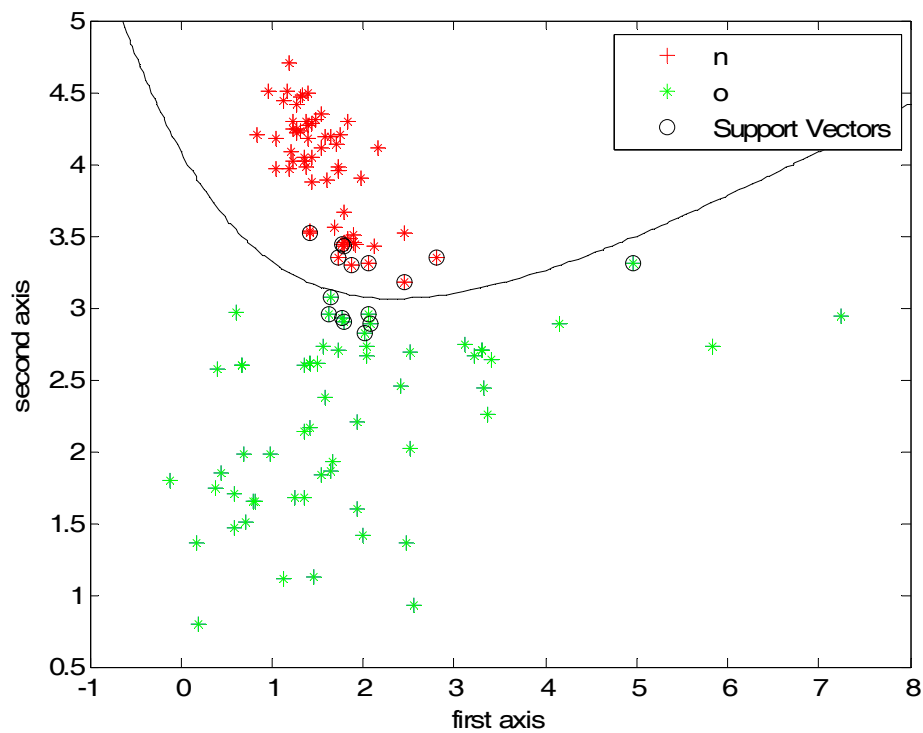


Figure 4.22 : Classification par SVM avec Bior3.3 et noyau polynomial du 2nd ordre.

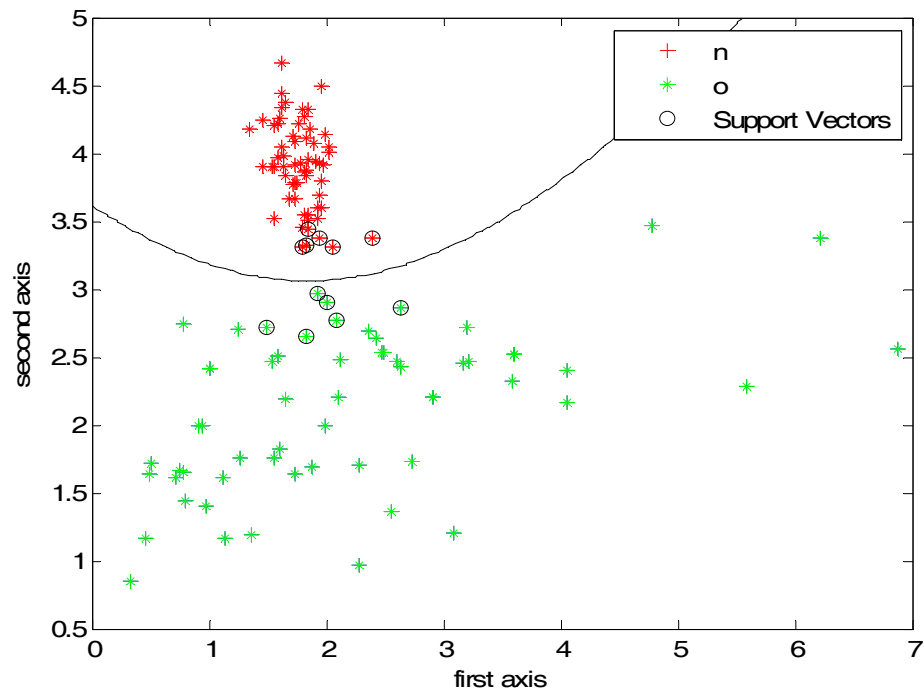


Figure 4.23 : Classification par SVM avec Bior5.5 et noyau polynomial du 2nd ordre.

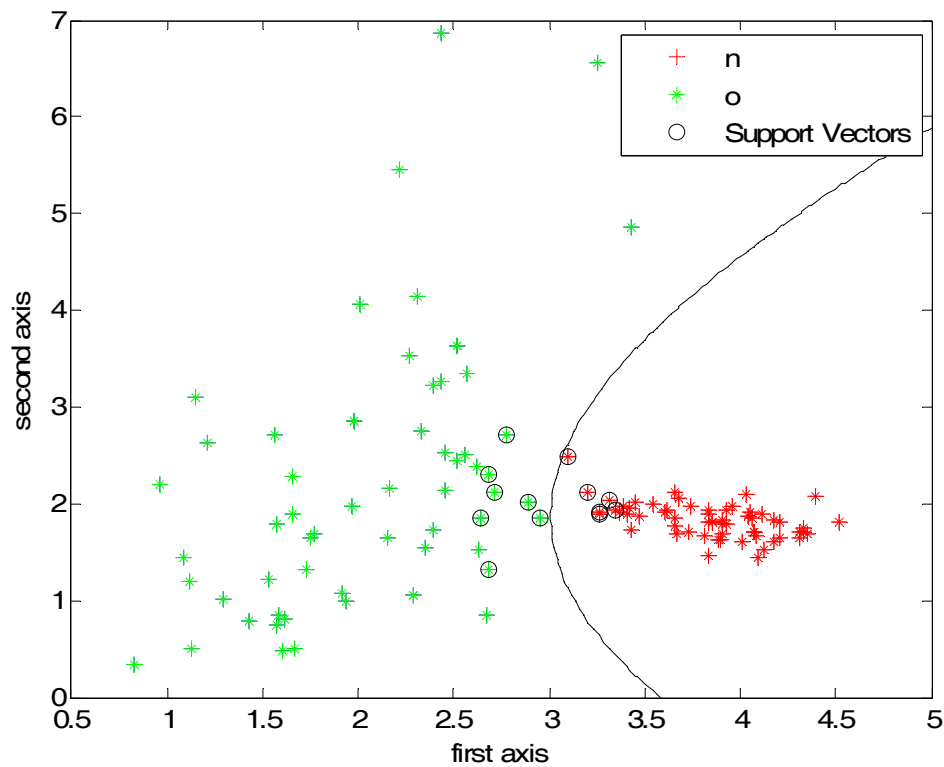


Figure 4.24 : Classification par SVM avec Sym6 et noyau polynomial du 2nd ordre.

Pour d'autres répartitions des données pour les étapes d'apprentissage et de test, nous appliquons la méthode de 10-fold cross validation, de la même manière que pour la première approche. Nous obtenons les résultats du tableau 4.10 qui donnent les pourcentages de la classification basée sur la combinaison DWT- ICA-SVM et utilisant un seul paramètre caractéristique (l'énergie de variance des coefficients de détail). La précision moyenne sur les dix estimations obtenues à partir des résultats des tests est considérée comme le critère d'évaluation des performances de l'algorithme de classification.

La figure 4.25 représente la distribution de TPR, TNR et TC à chaque itération de la validation croisée. Cette méthode a permis d'obtenir des valeurs optimales de 98.82%, 99.01% et 99,56% concernant, respectivement, la sensibilité, la spécificité, et la précision de classification de nos données d'enregistrement ECG . Le taux de 99.56% est justifié du moment que la base d'apprentissage varie d'une itération à l'autre.

Tableau4.10: Performances de l'approche temps-Fréquence

10-Fold	TPR%	TNR%	TC%
Test1	99.00	99.17	99.58
Test2	95.30	97.77	98.33
Test3	100	100	100
Test4	99.20	100	99.16
Test5	98,63	96.24	100
Test6	100	100	100
Test7	100	100	100
Test8	96.04	97.80	98,40
Test9	100	100	100
Test10	100	100	100

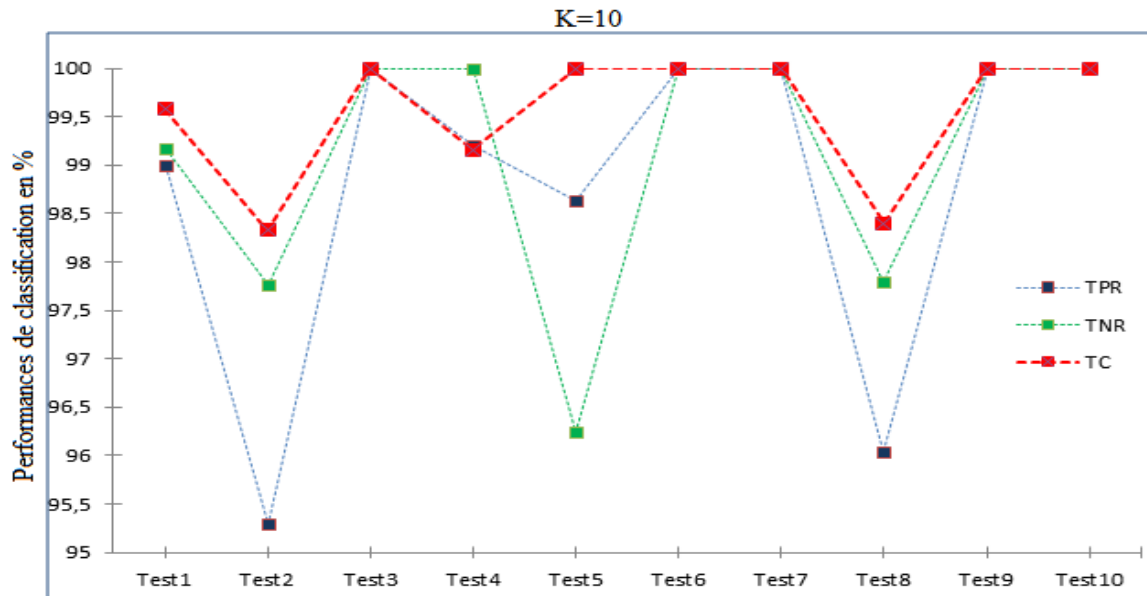


Figure 4.25: Performances de classification de l'approche DWT-ICA-SVM avec la validation croisée.

Le choix de la base de données est très important. Nous devons veiller, lors de sa conception, à ce que chaque enregistrement du cas pathologique comprenne un intervalle de temps où l'anomalie apparaît. Sinon, le taux de classification peut être affecté.

4.4.3 Temps de calcul des différentes étapes de décision

Notre première méthode RBF-GOFR-SVM a l'avantage d'exploiter la représentation temporelle du signal et nécessite un découpage du signal ECG en battements cardiaques. Notre contribution à cette approche est le choix des paramètres temporels et la bonne combinaison de méthodes RBF-GOFR-SVM pour atteindre la meilleure classification. Nous réalisons un bon taux de précision de 99,67%. Cette combinaison présente une nouvelle contribution et n'a pas encore été présenté dans la littérature. Dans l'approche temporelle, notre programme implémenté en Matlab R2011b (7.13.0.56.4), prend 15 s pour une heure d'enregistrement Holter. Sa mise en œuvre en langage assembleur pourrait réduire significativement ce temps de classification des enregistrements de battements normaux ou anormaux.

La deuxième méthode DWT-ICA-SVM ne nécessite aucun découpage du signal ECG en battements cardiaques mais a des contraintes sur la fréquence d'échantillonnage. Dans cette approche, nous montrons qu'une seule caractéristique statistique à savoir la variance d'énergie des coefficients de détails pour les cinq premiers niveaux de décomposition, sont suffisants pour représenter le signal ECG. Nous montrons qu'il est possible de réaliser une performance de classification appréciable c'est à dire une précision de 100%. Notre contribution a été de réduire

le nombre de caractéristiques pertinentes à une caractéristique statistique unique: la variance d'énergie, qui réduit le temps de traitement des signaux ECG.

Le temps de calcul alloué à l'ICA et SVM n'affecte en aucun cas le temps réel de traitement. Après la phase d'apprentissage, seule la matrice de transition doit être prise en compte pour l'ICA et l'équation de séparation pour le SVM.

A l'étape de décision chaque ECG passe par les étapes suivantes:

- décomposition par DWT et calcul de l'énergie de variance pour les coefficients de détail (paramètres statistiques).
- Multiplication par la matrice de transition (réduction d'espace par ICA).
- Décision après comparaison avec l'équation de séparation.

Avec un ordinateur personnel: TOSHIBA, Processor: Intel (R) Core (IM) I3 CPU M380@2.53 GHZ, Operating system: WIN32 bits et Matlab R2011b (7.13.0.56.4), le temps de calcul des différentes phases de décision est de 2.3s pour une heure d'enregistrements Holter. Il apparaît clairement que pour la mise en œuvre de la deuxième approche en temps réel, le temps est réduit de 7 fois par rapport à la première approche. Ce le temps peut encore être réduit grâce à d'autres mises en œuvre pour la deuxième approche, comme l'exploitation la technique de la transformée en ondelettes liftée (LWT), [116].

L'objectif dans nos travaux est de bien maîtriser les deux méthodes afin d'identifier leurs contraintes en temps réel. Certes, la base de données MIT database est importante dans le sens de la validation de nos méthodes, mais en pratique, nous devons nous préoccuper des différentes phases d'implémentation d'un système en temps réel.

Nous avons résumé dans le tableau 4.11 les principales caractéristiques pour les deux méthodes. Nous remarquons que la combinaison des algorithmes DWT, ICA et SVM est important pour atteindre le plus élevé score de classification de l'ECG.

Tableau4.11: Comparaison entre les deux approches

	Première approche	Seconde approche
Performance	99.67%	100%
Temps de calcul	15sec	2.3sec
Contraintes	Prétraitement et découpage	Choix de la fréquence d'échantillonnage

L'aspect comparatif des deux méthodes apparaît sur le temps de mise en œuvre avec un bon choix de la fréquence d'échantillonnage.

Nous avons jugé intéressant de comparer nos méthodes avec d'autres systèmes de reconnaissance présentés dans la littérature.

Quelques systèmes de reconnaissance de battements ECG représentatifs sont choisis pour cette comparaison: tels que la reconnaissance des ECG utilisant les réseaux de neurones hybrides flous (FHyb-HOSA), le modèle de réseau neuronal NNM (Neural Network Model), le neuro-flou (Neuro-Fuzzy), le système expert utilisant Kohonen et la décomposition en valeurs singulières SVD (SOM-SVD), la classification utilisant les algorithmes d'apprentissage par quantification vectorielle LVQ (Learning Vector Quantization) et autoregression AR MLP (MLP-LVQ) ainsi que Fourier et MLP (MLP-Fourier), [86].

Pour l'évaluation des résultats obtenus, le tableau 4.12 compare la précision de ces systèmes avec les nôtres. Comme des nombres différents de types de battements ont été exploités dans ces systèmes, c'est la précision moyenne de classification qui est calculée. Les résultats montrent que les méthodes proposées donnent relativement une plus grande précision de classification que les autres systèmes. La même base de données MIT-BIH database est utilisée pour l'évaluation de la performance de tous les systèmes.

Tableau4.12 : Résultats de comparaison de quelques méthodes de classification des battements cardiaques.

Méthode	Nombre du type de battement	Précision en %
RBF-GOFR-SVM (notre)	4	99.67
DWT-ICA-SVM (notre)	4	100
PNN-DWT	6	99.65
F Hyb-HOSA	7	96.06
NNM	5	97.78
Neuro-Fuzzy	4	98
SOM-SVD	4	92.2
MLP-LVQ	2	96.8
MLP-Fourier	3	98

Les résultats montrent que nos méthodes proposées fournissent relativement une précision de classification plus élevée que les autres systèmes. Notre deuxième approche DWT-ICA-SVM donne la meilleure performance.

Conclusion

Le principal objectif de notre thèse a été de proposer des techniques d'analyse et de traitement des signaux ECG, basées sur des méthodes mathématiques robustes de modélisation pour l'extraction d'un nombre minimal de paramètres pertinents. Nous avons testé leur pertinence sur des ECG issus de la base de données MIT et validé leur capacité à caractériser les ECG sains des anormaux. La continuité de nos travaux étant l'exploitation de la modélisation du signal ECG et sa compression en vue d'une détection d'anomalie, nous avons fait le choix d'un classificateur pertinent pour notre application, en occurrence le "SVM". Le but était de présenter au moins un classificateur de détection d'anomalies, parmi ceux qui existent dans la littérature afin de montrer l'efficacité des méthodes que nous avons développées.

La première approche de classification des battements ECG est basée sur l'utilisation des paramètres temporels où les battements du signal ECG sont détectés puis modélisés par l'algorithme Generalized Orthogonal Forward Regression (GOFR). Les paramètres (A_i^* , σ_i^* , μ_i^*) obtenus pour chaque fonction Gaussienne sont utilisés comme caractéristiques. La SVM est entraînée pour la classification avec la fonction noyau RBF. L'avantage de cette méthode par rapport aux méthodes courantes est le fait que cette technique extrait uniquement les paramètres temporels. Elle ne nécessite pas de transformations temps- fréquence. Dans ce cas, pour n'importe quelle classification nous avons la possibilité d'opérer en temps réel à partir d'une extraction en ligne des informations et donc d'établir une aide au diagnostic médical.

La seconde approche repose sur la transformée en ondelettes discrète DWT dont la décomposition en bandes de fréquences recouvre les caractéristiques fréquentielles des différentes ondes de l'ECG (P, QRS, T). L'exploitation des cinq premiers niveaux de décomposition a permis l'extraction des coefficients de détails pertinents de chaque signal ECG. Nous avons montré que le paramètre variance d'énergie des coefficients de détails était suffisant pour caractériser les signaux biomédicaux ECG et donc de réaliser une aide au diagnostic. L'ICA comme outil de réduction d'espace de mesure s'est avérée judicieux combinée au classificateur SVM. Le choix de la DWT de départ avait son importance afin de réaliser le meilleur score de classification.

La combinaison des paramètres temporeux issus de la modélisation par RBF et des paramètres fréquentiels issus de la décomposition par DWT avec l'algorithme de classification SVM a permis d'atteindre des taux de reconnaissance très compétitifs en termes de temps et taux de classification.

Conclusion générale

Afin de contribuer à la conception d'un système de diagnostic intelligent pour la détection des anomalies cardiaques des signaux ECG (électrocardiogramme), nous avons développé deux approches pour extraire les paramètres les plus pertinents qui peuvent le mieux classer les ECG en deux catégories : normaux ou anormaux.

Dans la première approche qui est une méthode temporelle, les battements du signal ECG sont détectés puis modélisés par l'algorithme combinant le réseau neuronal RBF (Radial Basis Function) avec Generalized Orthogonal Forward Regression (GOFR). Les ondes caractéristiques de chaque battement cardiaque sont modélisées par cinq gaussiennes avec une erreur de 10^{-5} . Les paramètres (A_i^* , σ_i^* , μ_i^*) obtenus pour chaque fonction Gaussienne sont utilisés comme caractéristiques pertinentes.

La seconde approche qui repose sur la décomposition temps-fréquence en utilisant la transformée en ondelettes discrète DWT (Discrete Wavelet Transform), a prouvé que le paramètre variance d'énergie des coefficients de détails était suffisant pour caractériser les signaux biomédicaux ECG. En effet l'information énergétique du signal est concentrée dans les coefficients de détails à plus de 90%. Le choix de la DWT de départ avait son importance afin de réaliser le meilleur score de classification. Les résultats obtenus montrent que le fait d'augmenter le nombre de paramètres représentant le signal n'améliore pas forcément la classification et ce, en comparaison avec la littérature. Autrement dit le fait de représenter le signal par un seul vecteur de cinq variances de niveaux de décomposition est largement suffisant s'il est suivi par une méthode d'extraction des composantes indépendantes.

Pour les deux approches, les données obtenues sont soumises à la même méthode de classification basée sur l'ACI comme outil judicieux de réduction d'espace de mesure et l'algorithme de classification SVM (Support Vector Machine). Le taux de reconnaissance dépend surtout du choix de la base de données des signaux anormaux qui est la MIT-BIH dans notre cas. Dans la conception de notre banque d'apprentissage, il est important de travailler étroitement avec les cardiologues, afin de séparer les battements normaux des anormaux. Il serait judicieux, dans des travaux futurs d'élargir la base de données pour englober plusieurs autres types de battements et augmenter ainsi la qualité de l'apprentissage.

Les résultats obtenus ont montré un bon ajustement des paramètres pertinents pour chaque approche et ont révélé la combinaison la plus efficace "traitement-classification". Nous avons comparé les deux approches et identifié certaines contraintes de mise en œuvre,

tels que la fréquence d'échantillonnage pour la méthode temps-fréquence et de prétraitement pour la méthode temporelle. La combinaison des algorithmes : DWT, ICA et SVM est déterminante pour la réalisation de la meilleure classification des signaux biomédicaux ECG lorsque le signal ECG à tester est de même fréquence d'échantillonnage que celle de la banque d'apprentissage. Nous avons atteint des taux de classification jusqu'à 100% avec une réduction du nombre de paramètres par la considération du seul paramètre statistique : la variance d'énergie des coefficients de détails des 5 premiers niveaux de décompositions de la DWT. Son temps d'implémentation est beaucoup plus réduit que pour la première approche. Ces résultats sont importants dans le sens où il ressort la possibilité de conception des systèmes intelligents pour l'aide au diagnostic et la surveillance des malades cardiaques en temps réel.

Suite à nos travaux d'analyse, de traitement et de classification des signaux biomédicaux ECG (électro cardiographiques enregistrés numériquement, nous envisageons en perspective la mise au point d'une implémentation hardware d'un système expert sous FPGA pour un diagnostic médical fiable et d'utilisation pratique, en cardiologie.

Pour l'implémentation hardware d'un prototype, nous allons exploiter les résultats de détection d'anomalies cardiaques, établis lors de l'implémentation software sous Matlab. Pour cela, nous aurons donc à maîtriser des outils informatiques tel que System Generator Xilinx (SGX) afin de matérialiser l'implémentation sous la carte FPGA, via VHDL.

Les étapes seraient donc les suivantes :

1. Réaliser un dispositif minimal d'acquisition des signaux biomédicaux en temps réel ECG adaptable aux entrées FPGA.
2. Rechercher des solutions aux problèmes de synchronisme entre le FPGA et les dispositifs d'acquisition, qui découlent du fait que notre prototype hardware soit conçu pour des applications en temps réel sur l'ensemble de la chaîne (acquisition-traitement- diagnostic).
3. Apprentissage et Maîtrise du logiciel System Generator Xilinx (SGX) qui permettrait de réaliser la transition entre les différentes méthodes de traitement sous Matlab, les convertir sous VHDL et passer ensuite à l'implémentation hardware sous FPGA.
4. Matérialisation de l'implémentation des différentes techniques de traitement du signal et de classification sous System Generator Xilinx (SGX) .
5. Implémentation sous FPGA (VHDL) c'est-à-dire qu'après avoir réalisé l'ensemble des programmes en VHDL avec SGX, passer à l'implémentation finale sous FPGA (Vertex5).

6. Optimisation des implémentations
7. Réalisation du prototype hardware pour la détection d'une anomalie cardiaque telle que l'arythmie cardiaque.

Parmi les travaux entrepris dans ce sens au sein de notre laboratoire, une thèse de doctorat intitulée : Implémentation hardware sous FPGA d'un système intelligent pour le monitoring des signaux biomédicaux, est entamée.

Nous projetons également de réaliser un système d'aide au diagnostic pour des pathologies cardiaques bien spécifiques et créer une banque de données algérienne : une pour les adultes et une autre pour les enfants dans le cadre de la médecine scolaire préventive.

Références

- [1] I. Guyon and A. Elisseeff. *An introduction to variable and feature selection*. Journal of Machine Learning Research 3, 1157-1182, 2003.
- [2] Y. H. Hu, S. Palreddy, and W. Tompkins. *A Patient Adaptable ECG Beat Classifier Using A Mixture Of Experts Approach*. IEEE Transactions on Biomedical Engineering vol. 44, pp. 891–900, 1997.
- [3] Costas Papaloukas, Dimitrios I. Fotiadis, Aristidis Likas, Lampros K. Michalis. *Automated Methods for Ischemia Detection in Long Duration ECGs*. Article in Cardiovascular reviews & reports · June 2003.
- [4] P. Tadejko, and W. Rakowski. *Mathematical Morphology Based ECG Feature Extraction for the Purpose of Heartbeat Classification*. 6th International Conference on Computer Information Systems and Industrial Management Applications, CISIM '07, pp. 322-327, 2007.
- [5] Felipe E. Olvera. *Electrocardiogram Waveform Feature Extraction Using the Matched Filter*. ECE 538/638: Statistical Signal Processing ece.pdx.edu/~ssp, 2006.
- [6] C. Alexakis, H. O. Nyongesa, R. Saatchi, N. D. Harris, C. Davies, C. Emery, R. H. Ireland, and S. R. Heller. *Feature Extraction and Classification of Electrocardiogram (ECG) Signals Related to Hypoglycaemia*. Conference on computers in Cardiology, pp. 537-540, IEEE, 2003.
- [7] Kuo-Kuang Jen, and Yean-Ren Hwang. *ECG Feature Extraction and Classification Using Cepstrum and Neural Networks*. Journal of Medical and Biological Engineering, vol. 28, no. 1, 2008.
- [8] A. Hariz. *Modélisation du signal Electrocardiogramme*. Mémoire de Magister, Laboratoire Génie-Biomédical, Département Instrumentation, FEI/USTHB, 2011, sous la direction de Pr M-D Kedir Talha.
- [9] A. B. Ramli, and P. A. Ahmad. *Correlation analysis for abnormal ECG signal features extraction*. 4th National Conference on Telecommunication Technology, 2003. NCTT 2003 Proceedings, pp. 232-237, 2003.
- [10] V. S. Chouhan, and S. S. Mehta. *Detection of QRS Complexes in 12-lead ECG using Adaptive Quantized Threshold*. IJCSNS International Journal of Computer Science and Network Security, vol. 8, no. 1, 2008.
- [11] C. Saritha, V. Sukanya, and Y. Narasimha Murthy. *ECG Signal Analysis Using Wavelet Transforms*. Bulgarian Journal of Physics, vol. 35, pp. 68-77, 2008.
- [12] H. Khorrami and M. Moavenian. *A Comparative Study of DWT, CWT and DCT in ECG Arrhythmias Classification*. Elsevier Expert Systems with Applications, February 2010, pp 5751-5757.
- [13] A. Mousa and R. Saleem. *Using Reduced Interference Distribution to Analyse abnormal Cardiac Signal*. Journal of Electrical Engineering, Vol 62, N°3, pp 168-172, 2011.
- [14] E. D. Ubeyli. *Time-Varying Biomedical Signals Analysis with Multiclass SVM employing Lyapunov Exponents*. Digital Signal Processing Archive, Volume 18 Issue 4, July, 2008, pp 646-656.
- [15] O. Faust, R. Acharya, S. M. Krishnan and L. C. Min. *Analysis of Cardiac Signals using spacial Filling index and Time Frequency Domain*. Biomedical Engineering Online 3, September 2004

- [16] Mazhar B. Tayel, and Mohamed E. El-Bouridy. *ECG Images Classification Using Feature Extraction Based On Wavelet Transformation And Neural Network*. ICGST, International Conference on AIML, June 2006.
- [17] Nait-Hamoud, A. Moussaoui. *Two Novel Methods for Multiclass ECG Arrhythmias Classification based on PCA, Fuzzy SVM and unbalanced Clustering*. International Conference on Machine and Web Intelligence (ICMWI), October 2010, pp 140-145.
- [18] Chou, K.-T. & Yu, S.-N. (2007). *Classifying ECG beats using ICA features and probabilistic neural network*. In: Magjarevic R., Nagel J.H. (eds) World Congress on Medical Physics and Biomedical Engineering 2006. IFMBE Proceedings, vol 14. Springer, Berlin, Heidelberg
- [19] G.G.Herrero, A.Gotchev, I.Christov, & K. Egiazarian. (2005). *Feature extraction for heartbeat classification using independent component analysis and matching pursuits*. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, March 18-23, 2005, pp.iv/725–iv/728, Philadelphia, PA, USA.
- [20] Ye, C., Coimbra, M. T. & Vijaya Kumar, B. V. K. (2010). *Arrhythmia detection and classification using morphological and dynamic features of ECG signals*. Proceedings of 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1918–1921, , Buenos, Aires, Argentina, August 31-September 4, 2010.
- [21] M. R. Homaeinezhad, E. Tavakkoli, M. Habibi, S. A. Atyabi, A. Ghifari. *Combination of different Classifiers for Cardia Arrhythmia Recognition*. World Academy of Science, Engineering and Technology , 2011.
- [22] J. R. Hampton, F.Jan. *L'ECG facile''*, 8 ième édition, ELSEVIER / MASSON, Collection : Facile et bio,Avril 2015.
- [23] A. R. Houghton and D. Gray. *Maîtriser l'ECG de la théorie à la Clinique*. Editeur ELSEVIER, Date de parution :22 avril 2011, ISBN 13 :9782294711602.
- [24] D.Hebert. *Lire un ECG*. Formation sur le langage et la lecture d'un ECG destinée aux ambulanciers. Février 2006.
- [25] A.Micheau, D.Hoa, M.Arnoult . *Cardiologie vasculaire*. Vernazobres-Greco, pp 40-59. 1977.
- [26] W. Einthoven. *The galvanometric registration of the human electrocardiogram* likewise a review of the use of capillary-electrometer in physiology, in Cardiac classics: Willius FW-CV Mosby, St Louis, 1941.
- [27] C. Guerot et Y. Grosgeat. *L'électrocardiogramme : Savoir l'interpréter'*. pp.11-16 ;pp.20-23 ; pp.27-28. Paris 1995.
- [28] J. Adamec · R. Adamec. *ECG Holter Guide to Electrocardiographic Interpretation* .Foreword I by Prof. Lukas Kappenberger. Foreword II by Prof. Philippe Coumel . DOI: 10.1007/978-0-387-78187-7 Library of Congress. Springer. 2008.

- [29] F. Fillette, G. Fontaine and B. Tardieu. *L'essentiel sur l'enregistrement Holter de l'ECG*. Masson, 1983.
- [30] A. Cabasson. *Estimation et analyse des intervalles cardiaques*. Thèse de Doctorat, Université de Nice - Sophia Antipolis, 2008.
- [31] Sara Wong. *Segmentation de l'intervalle RT et description par analyse factorielle de la variabilité de la fréquence cardiaque et de la repolarisation ventriculaire*. Thèse de Doctorat, Université de Rennes1, Juin 2004.
- [32] O. Fokapu and J.P. Girard. Evolution temporelle et fréquentielle de l'ECG : analyse battement par battement. ITBM, pages 102-112, 1994.
- [33] S. Rerbal, M. Benabdellah, N.Habibes, A.Meziane Tani, A. Nemmiche. *Analyse Spectro-Temporelle du Signal Electro-cardiographique Dans le cas de deux Cardiopathies*. Université Abou Bekr Belkaid –Tlemcen.
- [34] J. W. Hurst. *Ventricular Electrocardiography*. Lippincot Williams &Milkins Publishers, 1990.
- [35] J. Morganroth, I . Gussak. *Cardiac safety of noncardiac drugs*. practical guidelines for clinica . p139, 2005.
- [36] J.C.KAHN. *Trouble de rythme ventriculaire*. CHVII D,PP.1-8.
- [37] P. Defaye et G. Vanzetto. *Extrasystole ventriculaire*. 336a, Octobre 2002 , Mise à jour janvier 2005.
- [38] P. Naken. *A metric for line segments*. IEEE Trans. Pattern Anal. Machine Intell., vol. 15, pp. 1312-1318, 1993.
- [39] T. Pavlidis and S. Horowitz. *Segmentation of plane curves*. IEEE Trans. Computers , vol. 23, pp. 860-870, 1974.
- [40] B.Ray and K. Ray.*An algorithm for polygonal approximation of digitized curves*. Pattern Recognition Lett., vol. 13, pp. 489-496, 1992.
- [41] A. Koski.*Modelling ECG Signals with Hidden Markov Models*. Artificial Intelligence in Medicine, vol. 8, pp. 453-471, 1996.
- [42] S. Mallat. *Une exploration des signaux en ondelettes*. Les Editions de l'Ecole Polytechnique ed, 2000.
- [43] B. Torr sani. *Analyse continue par ondelettes*. InterEditions / CNRS Editions ed: Savoirs Actuels, 1995.
- [44] Poularikas. *The Transforms and Applications Handbook*. IEEE Press ed: CRC Press, 1996.
- [45] L. Thoraval, G. Carrault and J. J. Bellanger. *Heart Signal Recognition by Hidden Markov Models: The ECG Case*. Methods of Information in Medicine, vol.33, pp. 10-14, 1994.
- [46] L. Senhadji, L. Thoraval and G. Carrault. *Continuous wavelet transform: ECG recognition based on phase and modulus representations and hidden Markov model*. Wavelet Applications in Medicine and Biology, pp. 439-463, 1996.

- [47] M. Bahoura, M. Hassani and M. Hubin. *DSP Implementation of Wavelet Transform for Real Time ECG Wave Forms Detection and Rate Analysis*. Computer Methods and Programs in Biomedicine, vol. 52, pp. 35-44, 1997.
- [48] M. L. Hilton. *Wavelet and Wavelet Packet Compression of Electrocardiograms*. IEEE Transactions on Biomedical Engineering, vol. 44, pp. 3944-3952, 1997.
- [49] S. M. Ahmed, A. Al-Sharouf and M. Abo-Zahhad. *ECG Data Compression using Optimal non-Orthogonal Wavelet*. Medical Engineering & Physics, vol. 22, pp. 39-46, 2000.
- [50] Broomhead, D.S., Lowe, D. (1988). *Multivariate functional interpolation and adaptive networks*. Complex Systems, vol 2, p.321-355.
- [51] J. Moody and C.J. Darken (1989). *Fast Learning in Network for Locally Tuned Processing Units*. Neural Computation, vol.1, p.281-294.
- [52] T. Poggio and F. Girosi (1990). *Network for approximation and learning*. Proc. IEEE, vol 78, no. 9, pp. 1481 -1497.
- [53] G. Dreyfus, M. Samuelides, M. Martinez, M. Gordon , et al. *Réseaux de Neurones*. Eyrolles ed, 2002.
- [54] M. Kedir-Talha, A. Hariz and S. Ould-Slimane. *Modelling of the beat of a cardiac signal by Gaussians*. Asilomar conference USA Nov 2010 DOI: 10.1109/ACSSC.2010.5757843
- [55] M. Kedir-Talha and S. Ould-Slimane. *Treatment of cardiac signal for a modeling by RBF*. ISABEL 2011 Fourth International Symposium on « Applied Sciences in Biomedical and Communication Technologies» October 26-29, 2011, Barcelona, Catalonia, Spain, ACM New York, NY, USA ©2011, ISBN: 978-1-4503-0913-4 doi>10.1145/2093698.2093757
- [56] M.D. Kedir-Talha and S. Ould-Slimane. *Neural Networks and SVM for Heartbeat Classification*. IEEE – 2012, 11th International Conference on Information Sciences”Signal Processing and their Applications (ISSPA 2012). 3-5 Juillet Montreal Canada.
- [57] R. Dubois, P. Maison-blanc, B. Quenet, and G. Dreyfus, “Automatic ECG Wave Extraction in Long-term Recordings using Gaussian mesa Function Models and Nonlinear Probability Estimators,” *Computer Methods and Programs in Biomedicine*, vol. 88, no 3, pp. 217-233, 2007.
- [58] S. Chen, S. A. Billings and W. Luo, "Orthogonal least squares methods and their application to non-linear system", *Int. J. Control*, vol. 50, pp. 1873-1896, 1967.
- [59] Y. Oussar and G. Dreyfus, "Initialization by selection for wavelet network training", *Neurocomputing*, vol. 34, pp. 131-143, 2000.
- [60] <http://www.physionet.org/physiobank/database/mitdb/>
- [61] S. Ould Slimane and M. Kedir-Talha. *Prétraitement du signal cardiaque en vue d'une modélisation*. Communication 25-26 Octobre 2010. JLIN'S2010.FEI-USTHB.
- [62] J. Stephenson. *Detection of Isoelectric Baseline and High Frequency Noise within Electrocardiographic Signal*. ESPCI, Laboratoire d'Electronique 2000.

- [63] J. Pan and W. J. Tompkins. *A Real-Time QRS Detection Algorithm*. IEEE Transactions on Biomedical Engineering, vol. BME 32, pp. 246, 1985.
- [64] C. Nugent, and J. Lopez. *Prediction Models in the Design of Neural Network Based ECG Classifiers: A neural network and genetic programming approach*. BMC Med Inform Decis Mak. 2002.
- [65] R.Polikar. *The wavelet tutorial*. The engineer's ultimate guide to wavelet analysis. Juin 1996.
- [66] Y. Meyer. *Les ondelettes, algorithmes et applications*. Edition Armand Colin.1993/1994.
- [67] S. Mallat. *A wavelet tour of signal processing*. Academic Press, 1998.
- [68] Y. Oussar. *Réseau d'ondelettes et réseaux de neurones par la modélisation statique et dynamique de processus*. Thèse de doctorat. Université Pierre et Marie Curie.1998.
- [69] M.Holschneider. *Localization properties of wavelet transforms*. J. Math. Phys, 34(7), 3227–3244, 1993.
- [70] M.Abo-Zahhad, A. F.Al-Ajlouni, S. M.Ahmed, R.J.Schilling. *A new algorithm for the compression of ECG signals based on mother wavelet parameterization and best-threshold levels selection*.Digital signal processing. Elsevier.Volume 23, Issue 3, May 2013, Pp 1002-1011.
- [71] N.Thi Hoàng Lan . *Étude de la méthode de la transformation en ondelette et l'application à la compression des images*, 2005.
- [72] D. O.Walter. *Spectral analysis for electroencephalograms : mathematical determination of neurophysiological relationships from records of limited duration*. Exp. Neurol.,8:155–181, 1963.
- [73] A. Ben attalah, M. A. Ben ayed, M. Loulou, N. Masmoudi. *Implémentation électronique de la transformer en ondelette discrète unidimensionnelle*. Conférence SETIT 2003, 17 Mars-21 Mars 2003 Sousse, Tunisie.
- [74] I.Laraba et S.Kouici. *Compression par ondelettes d'un signal ECG*. Projet de Fin d'Etudes d'Ingénieur d'Etat en Electronique, Option : Contrôle , FEI/USTHB,2008.
- [75] ECG Databases. 15-May-2008. <http://www.physionet.org/physiobank/database/>.
- [76] I. Daubechies. *Ten lectures on wavelets* . CBMS, SIAM, 1994.
- [77] M. L. Hilton. *Wavelet and Wavelet Packet Compression of Electrocardiograms*. IEEE Transactions on Biomedical Engineering, vol. 44, no. 5, pp. 394-402, 1997.
- [78] B. R. Shankara et I. S. N. Murthy. *ECG Data Compression Using Fourier Descriptors*. IEEE Transactions on Biomedical Engineering, vol. BME-33, no. 4, pp. 428-434, 1986.
- [79] Redha Benzid. *Ondelettes et statistiques d'ordre supérieur appliquées aux signaux uni et bidimensionnels*. Thèse de Doctorat en Electronique-Option Contrôle,2005. Université de Batna.
- [80] Jérôme Carette. *Les ondelettes et la compression d'images*. Département d'informatique et de mathématique Université du Québec à Chicoutimi,décembre 2000.
- [81] R. Ardon, F. Perret, C. Roche. *Débruitage par seuillage de coefficients d'ondelettes*. ENS cachain Paris.2001.

- [82] Cyprien Dumortier. *La compression d'image par ondelette*. Ecole polytechnique de Montréal.
- [83] M. Kedir-Talha, S.Ould-Slimane and A. Bousbia-Salah. *Compression of biomedical signals ECG and EEG*. WOTIC 2011, International workshop on Information Technologies and communication. October 13-15,2011. Casablanca, Morocco.
- [84] M.Talha-Kedir, I.L, S.K, S.Ould Slimane. *Comparison between the orthogonal Wavelet and the biorthogonal, for the transmission of a cardiac signal*. International conference on Systems and Signal Information, Mai 2009 Guelma.
- [85] Saliha Ould Slimane, Malika Kedir Talha, Zairi Hadjer and Slimane Mekaoui. *Comparison of time and frequency methods for cardiac anomalies recognition*. Journal of Biomedical Engineering - Applications Basis and Communications Vol (28) issue 2, April 2016 DOI: 0.4015/S1016237216500101.
- [86] Kedir-Talha MD, Ould-Slimane S. *ECG and EEG Biosignals Classification by Hybrid DWT-ICA-SVM*. International Journal of Computational Bioscience, Acta press ,Edition, 4:1, pp. 11–17, 2013. DOI: 10.2316/J.2013.210-1063
- [87] Pearson, K. 1901. *On lines and planes of closest fit to systems of points in space*. Philosophical Magazine 2:559-572. <http://pbil.univ-lyon1.fr/R/pearson1901.pdf>
- [88] H. Hotelling. *Analysis of a complex of statistical variables into principal components*. Journal of Educational Psychology, 1933.Publisher: American Psychological Association (APA)DOI: 10.1037/h0071325
- [89] Pierre-Louis Gonzalez. *L'analyse en composante principale*.Rappel de cours.
- [90] Chihoon Lee. *Principal Component Analysis*. Computing Science Dept. University of Alberta. 15 September 2004.
- [91] Gilbert Saporta. *Probabilités, analyse des données et statistique*. Editions TECHNIP. 2006.
- [92] Hervé Abdi andLynne J. Williams. *Principal component analysis*. Wiley Interdisciplinary Reviews: Computational Statistics. 2010.
- [93] Nathalie Villa-Vialaneix.Analyse en composantes principales. M2 Statistique et Econométrie. Data Mining 1. Année 2008/2009.
- [94] Aapo Hyvärinen and Erkki Oja. *Independent Component Analysis Algorithms and Applications*. Neural Networks Research Centre Helsinki University of Technology. 2000.
- [95] Le Borgne Hervé. *Analyse de Scènes Naturelles par Composantes Indépendantes*. Thèse de doctorat de l'Institut National Polytechnique de Grenoble. Janvier 2004.
- [96] Gobinet Cyril. *Application des technique de la séparation de sources à la spectroscopie RAMAN* . Thèse de doctorat de l'université Reims Champagne-Ardenne.2006.
- [97] Hyvärinen, A. (1999). *Fast and robust fixed-point algorithms for independent component analysis*. IEEE Transactions on Neural Networks, vol. 10, no 3, p. 626-634.
- [98] Jutten, C. (1987). *Calcul neuromimétique et traitement du signal, analyse en composantes indépendantes*. Thèse de doctorat, Université de Grenoble, Grenoble, 200 p.

- [99] Weems A, Harding M, Choi A. *Classification of the ECG Signal Using Artificial Neural Network*. Proceedings of the 3rd International Conference on Intelligent Technologies and Engineering Systems (ICITES2014) pp 545-555.
- [100] Gaurav Kumar Jaiswal and Ranbir Paul. *Artificial neural network for ecg classification*. Recent Research in Science and Technology 2014, 6(1): 36-38 ISSN: 2076-5061.
- [101] Jaswantsing L. Rajput, Mrs. C. S. Khandelwal. *Classification of ECG Abnormalities using ANN*. International Journal of Engineering Research & Technology, Vol.1, Issue 5 (July 2012), e-ISSN: 2278-0181.
- [102] Stalin Subbiah, Rajkumar Patro, Subbuthai P. *Feature Extraction and Classification for ECG Signal Processing based on Artificial Neural Network and Machine Learning Approach*. International Conference on Inter Disciplinary Research in Engineering and Technology (2015):50-57.Print.
- [103] Saini I, Singh D, Khosla A. *QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases*. J ADV RES. 2013 JUL; 4(4): 331–344.PUBLISHED ONLINE 2012 JUL 6. DOI: 10.1016/J.JARE.2012.05.007.
- [104] Saini I, Singh D, Khosla A. *K-nearest neighbour-based algorithm for P- and T-waves detection and delineation*. J Med Eng Technol. 2014 Apr;38(3):115-24. doi: 10.3109/03091902.2014.882424. Epub 2014 Feb 10.
- [105] Korurek M, Dogan B(2010). *ECG beat classification using particleswarm optimization and radial basis function neural network*. Expert Sys Appl 33;7563-7569.
- [106] Talbi M.L. *Analyse et traitement du signal électrographique (ECG)*. Thèse de Doctorat en Sciences en Electronique. Université MENTOURI de CONSTANTINE (2011).
- [107] Mehrdad Javadi, Reza Ebrahimpour, Atena Sajedin, Soheil Faridi, and Shokoufeh Zakernejad. *Improving ECG Classification Accuracy Using an Ensemble of Neural Network Modules*. PLoS One. 2011; 6(10): e24386. doi: 10.1371/journal.pone.0024386.
- [108] Moavenian M, Khorrami H. *A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification*. Expert Systems with Applications 37 (2010) 3088–3093.
- [109] V. Vapnik. *The nature of statistical learning theory*. Springer-Verlag ed, Springer, New York, 1995.
- [110] Burges CJC. *A tutorial on support vector machines for pattern recognition*. Kluwer Academic Publishers, Boston, 1998.
- [111] Bottou L. and Chih-jen L. *Support Vector Machine Solvers in Large Scale Kernel Machines*. MIT Press. 2007.
- [112] Mercier L. *Les machines à vecteurs support pour la classification en imagerie hyper spectrale : implémentation et mise en œuvre*. Février 2010.
- [113] Nassima Hamdad. *Transformée de Huang-Hilbert: Application à la détection des défauts*. Thèse de doctorat, Université Mouloud Mammeri, Tizi-Ouzou (UMMTO), Algerie, 2013.

- [114] Mounia Hendel, Abdelkader Benyettou et Hiba Khelil. *Classification des Arythmies Cardiaques par les Réseau de Neurones Artificiels*. Université des Sciences et de Technologie d'Oran Département d'informatique, laboratoire SIMPA El M'nouar BP 1505, Oran Algérie.
- [115] Abdulkadir S. *Multiclass least-square support vector machines for analogmodulation classification*. Expert System with Applications, Vol. 36, pp. 6681–6685. 2009.
- [116] Kedir-Talha MD, Ait Amer MA, *The lifted wavelet transform for encephalic signal compression*, 36th International Conference on Telecommunications and Signal Processing, IEEE, TSP, pp. 541–544, Italy, 2013.

Les filtres miroirs conjugués

Ils sont caractérisés par l'utilisation de filtre h vérifiant :

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega + \pi)|^2 = 2$$

où $h(\omega)$ est un polynôme trigonométrique.

Le déphasage de π revient à changer le signe d'un coefficient sur deux. On cherche de plus à fixer le nombre de zéro de la fonction de transfert $h(n)$ dans le but d'annuler les moments de l'ondelette.

Afin de calculer $|\hat{h}(\omega)|^2$ on utilise le théorème de factorisation spectrale :

1. Théorème de Riesz :

Soit A un polynôme trigonométrique positif invariant sous l'action de $\omega \rightarrow -\omega$. A est nécessairement de la forme :

$$A(\omega) = \sum_{m=0}^M a_m \cos(m\omega) \text{ avec } a_m \in \mathbb{R}$$

Alors il existe un polynôme trigonométrique B d'ordre M , c-à-d :

$$B(\omega) = \sum_{m=0}^M b_m e^{m\omega} \text{ avec } b_m \in \mathbb{R}$$

$$\text{tel que } |B(\omega)|^2 = A(\omega)$$

Cela permet de trouver h .

2. Les filtres RIF :

Le filtre h détermine entièrement la fonction d'échelle et la plupart de ses propriétés (nullité des moments, support, régularité et symétrie). En particulier, la fonction d'échelle est à support compact ssi h a un nombre finie de coefficients nuls. On dit alors que h est un filtre RIF : à réponse impulsionnelle finie (FIR filter : finite impulse response).

Condition imposée au filtre RIF pour assurer la reconstitution du signal :

Considérons le cas d'un filtrage d'un signal $x(t)$ par un filtre $h(t)$. Il est représenté dans le domaine temporel par le produit de convolution suivant :

$$y(t) = h(t) \otimes x(t)$$

soit dans le domaine fréquentiel :

$$Y(e^{j\omega T_s}) = H(e^{j\omega T_s})X(e^{j\omega T_s}) \text{ avec } H(e^{j\omega T_s}) = A(e^{j\omega T_s})e^{j\varphi(f)}$$

$$Y_n = T_s \int_{-1/2T_s}^{+1/2T_s} A(e^{j\omega T_s})e^{j\varphi(f)}X(e^{j\omega T_s})e^{j2\pi f n T_s}df$$

Pour un filtrage au voisinage d'une fréquence η , la phase peut être approximée par un développement limité au premier ordre :

$$\begin{aligned} \varphi(f) &= \varphi(\eta) + 2\pi(f - \eta) \cdot \left. \frac{d\varphi}{df} \right|_{f=\eta} \Rightarrow y_n \\ &= T_s e^{j\varphi(\eta)} \int_{-1/2T_s}^{+1/2T_s} A(e^{j\omega T_s})X(e^{j\omega T_s})e^{-j2\pi(f-\eta)\tau(\eta)}e^{j2\pi f n T_s}df \end{aligned}$$

Le filtre introduit donc un retard :

$$\tau(\eta) = \left. \frac{d\varphi}{df} \right|_{f=\eta}$$

avec τ le temps de groupe (ou retard de groupe) pour les fréquences voisines de η .

Le signal $x(t)$ a un spectre constitué de plusieurs fréquences (comme dans le cas de notre étude), les différents groupes de fréquences seront transmis (décomposé) avec des retards différents.

Cela entraînera une distorsion lors de la reconstitution de y_n .

Pour pallier à cet inconvénient, le temps de groupe τ doit être indépendant de la fréquence, soit satisfaire la condition :

$$\tau(\eta) = \left. \frac{d\varphi}{df} \right|_{f=\eta} = cste \Rightarrow \varphi(f) = \varphi_0 - 2\pi f \tau.$$

Cela signifie qu'on doit avoir une phase linéaire.

a. Condition d'obtention d'une phase linéaire :

La phase linéaire peut être obtenue si les deux conditions suivantes sont satisfaites :

_ Le filtre est à phase linéaire,

_les coefficients h_n du filtre sont réels.

Les coefficients du filtre ont pour expression :

$$h_h = T_s e^{j\varphi_0} \int_{-1/2T_s}^{+1/2T_s} A(e^{j\omega T_s}) e^{-j2\pi f t} e^{j2\pi f n T_s} df$$

où $A(e^{j\omega T_s})$ est le module de $H(e^{j\omega T_s})$ (le module est de type réel).

Si on exprime $h(t+\tau)$, les coefficients s'écrivent :

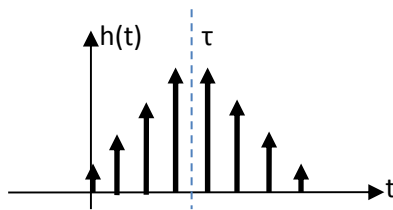
$$\begin{aligned} h_{n+\frac{\tau}{T_s}} &= T_s e^{j\varphi_0} \int_{-1/2T_s}^{+1/2T_s} A(e^{j\omega T_s}) e^{j2\pi f n T_s} df \\ &= T_s e^{j\varphi_0} \left[e^{j\varphi_0} \int_{-1/2T_s}^{+1/2T_s} A(e^{j\omega T_s}) \cos(2\pi f n T_s) df \right] + \\ T_s &\left[e^{j\varphi_0} \int_{1/2T_s}^{+1/2T_s} A(e^{j\omega T_s}) \sin(2\pi f n T_s) df \right] \end{aligned}$$

La phase linéaire $\Phi(f) = \Phi_0 - 2\pi f \tau$ s'obtient en pratique en considérant deux cas particuliers en prenant $\Phi_0=0$ et $\Phi_0=\frac{\pi}{2}$

a. Premier cas : $\Phi_0=0$:

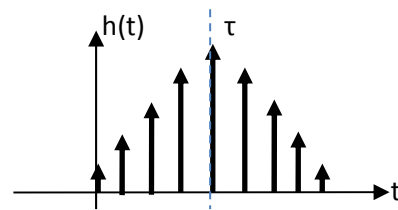
$A(e^{j\omega T_s})$ étant réel, pour que $h(t+\tau)$ le soit aussi, le second terme de la formule précédente doit s'annuler et le premier terme doit être pair par rapport à t : $h(t+\tau)$.

On en déduit que les échantillons de la réponse impulsionnelle sont réels et symétriques par rapport à τ/T_s , la réponse impulsionnelle est réelle et paire. Cela est obtenu dans les deux cas de figure suivants :



Nombre pair d'échantillons

$$N=8 ; \tau=T_s(N-1)/2$$



Nombre impair d'échantillons

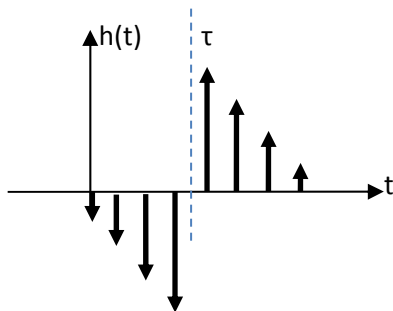
$$N=9 ; \tau=T_s(N-1)/2$$

b. **Deuxième cas** : $\Phi_0=\pi/2$:

Sachant que $e^{-j\frac{\pi}{2}} = -j$ et $A(e^{j\omega T_s})$ est réel, pour que $h(t+\tau)$ le soit également, le premier terme de la formule doit s'annuler et le second terme doit être impair par rapport à t :

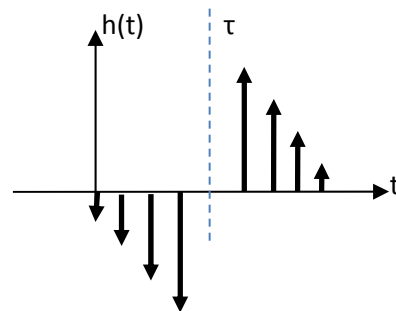
$$h(t+\tau)=h(-t+\tau)$$

Les échantillons de la réponse impulsionnelle sont donc réels et antisymétrique par rapport à τ/T_s : la réponse impulsionnelle est réelle et impair. Cela est obtenu dans les deux cas de figure suivants :



Nombre pair d'échantillons

$$N=8 ; \tau=T_s(N-1)/2$$



Nombre impair d'échantillons

$$N=9 ; \tau=T_s(N-1)/2 ; h_0=0$$

Pour les deux cas particuliers étudiés, la parité ou l'imparité confère une symétrie pair ou impair à la réponse impulsionnelle, ce qui permet en pratique de réduire de moitié le nombre de coefficients stockés en mémoire. Cela est intéressant puisque ces filtres RIF peuvent avoir plusieurs centaines de coefficients.

I. Caractéristiques des différentes ondelettes

En se basant sur les propriétés des ondelettes et les possibilités d'analyse qu'elles permettent de réaliser, on peut classer les ondelettes en cinq groupes :

ONDELETTES	PROPRIETES ET ANALYSE POSSIBLE
Ondelettes brutes (ondelettes : Gaussienne, Chapeau mexicain)	_ Φ n'existe pas et l'analyse n'est pas orthogonale. _ Ψ a une expression explicite mais pas à support compact. _ La propriété de reconstitution n'est pas assuré. _ La décomposition continue est possible. _ L'application d'algorithme rapide est impossible. _ La propriété de symétrie est assuré.
Ondelettes infiniment dérivable (ondelette de Meyer)	_ Φ existe et l'analyse est orthogonale. _ Φ et Ψ sont indéfiniment dérivables. _ Φ et Ψ ne sont pas à support compact. _ Les transformées continue et discrète sont possible mais sans filtre FIR. _ Les propriétés de symétrie et de régularité sont assurées. _ L'utilisation des algorithmes rapide est impossible.
Ondelettes orthogonales à support Compact (ondelettes : de Daubechies, Symlets et de Coiflets)	_ Φ existe et l'analyse est orthogonale. _ Φ et Ψ ne sont pas à support compact. _ Ψ à un nombre donné de moment nuls. _ La transformée continue est applicable. _ La transformée discrète est applicable en utilisant la WT. _ La régularité est médiocre. Propriétés spécifiques : Pour Daubechies : asymétrie. Pour Symlet : symétrie relative. Pour coiflet : symétrie relative. Ψ et Φ ont des moments nuls.
Pair d'ondelettes biorthogonales à support compact (ondelettes	_ Les fonctions Φ existent et l'analyse est biorthogonale. _ Φ et Ψ ont un support compact pour la décomposition et la reconstitution. _ Φ et Ψ ont une bonne régularité pour la reconstitution. _ Φ et Ψ ont des moments nuls pour la décomposition. _ La transformé continue est applicable.

biorthogonale B-spline)	_ La transformée discrète est applicable en utilisant la FWT. _ Les filtre FIR sont symétriques. _ Les fonctions Φ et Ψ ne sont pas orthogonales.
----------------------------	---

Ondelettes complexes (gaussiennes, de Morlet, de Shanon et B-spline en fréquence)	_ Φ n'existe pas et l'analyse n'est pas orthogonale. _ Ψ a une expression explicite mais pas à support compacte. _ La propriété de reconstitution n'est pas assurée. _ La décomposition complexe continue est possible. _ L'utilisation d'algorithme rapide est impossible.
---	---

II. Orthogonalité et biorthogonalité des ondelettes

1. L'orthogonalité

L'orthogonalité signifie que l'information capturée par une ondelette est totalement décorrélée de celle capturée par une autre. Cela permet de ne garder que l'information nécessaire et suffisante afin d'assurer la réversibilité. Une ondelette mère orthogonale et sa fonction d'échelle permettent d'effectuer une analyse multi-résolution orthogonale : les espaces de détail et d'approximation sont alors orthogonaux, c'est-à-dire que la projection des vecteurs de la base de l'un des espaces sur l'autre donne zéro. Chaque ondelette et sa fonction d'échelle associée sont également orthogonales [79].

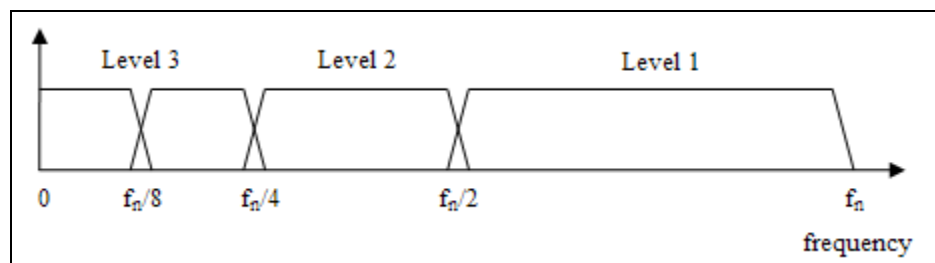


Figure 1: Spectre des ondelettes orthogonales

La fonction d'ondelette Ψ représente un filtre passe haut, qui définira les détails, sa fonction correspondante aura pour expression :

$$\Psi(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} g_n[n] \Phi(2t - n) \quad (1)$$

Les approximations sont définies par la fonction d'échelle Φ qui représente un filtre passe bas, sa fonction correspondante est :

$$\Phi(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} h_n[n] \Phi(2t - n) \quad (2)$$

Un exemple d'ondelette orthogonale l'ondelette de Haar (figure 2) permettant le calcul des deux types de transformée en ondelette, continue et discrète. Elle est définie par :

$$\Psi(x) = \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2}[\\ -1 & \text{si } x \in [\frac{1}{2}, 1[\\ 0 & \text{ailleurs} \end{cases}$$

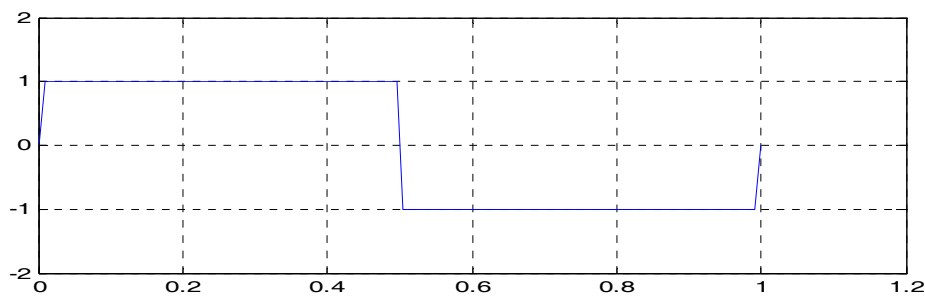


Figure 2: Tracé de la fonction d'ondelette de Haar

L'implémentation de la transformée en ondelette orthogonale se fait par application d'un algorithme rapide à base de bancs de filtre. On utilise pour cela une paire de filtres (**h** et **g**) miroirs conjugués à réponse impulsionnelle finie (**RIF**), **h** est de type passe bas et **g** de type passe haut .

2. L'analyse multirésolution orthogonale

Les filtres **h** et **g**, qui déterminent un banc de filtres **RIF** à reconstruction parfaite, permettent d'effectuer des sous-échantillonnages du signal par suppression d'un terme sur deux au cours de la décomposition, donc pas de redondance (figure 3), et des sur-échantillonnages par insertion d'un zéro tous les deux coefficients au cours de la reconstitution (figure 4).

Par conséquent, on aura un meilleur facteur de compression et le signal d'origine pourra être reconstruit.

Par contre, si les fonctions $\phi(t)$ et $\psi(t)$ ne sont pas orthogonales, deux conséquences importantes en découlent:

- on aura à compresser plus d'informations sur le signal qu'il n'en faut,
- la restitution du signal d'origine à partir des transformées sera impossible.

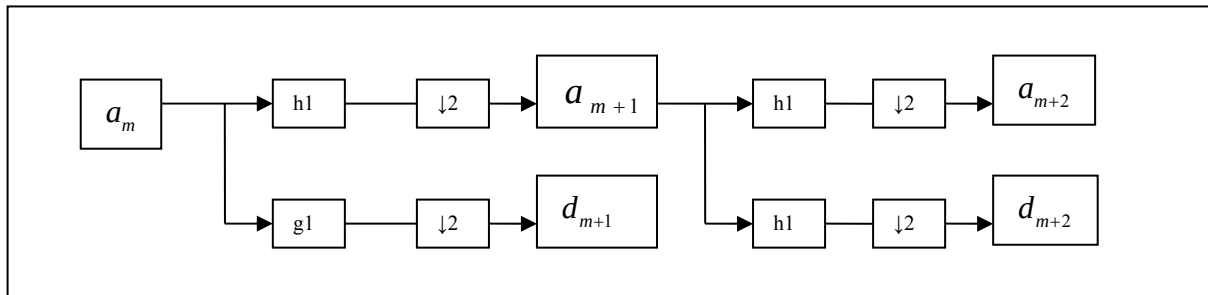


Figure 3 : Décomposition à base d'ondelette avec filtres RIF.

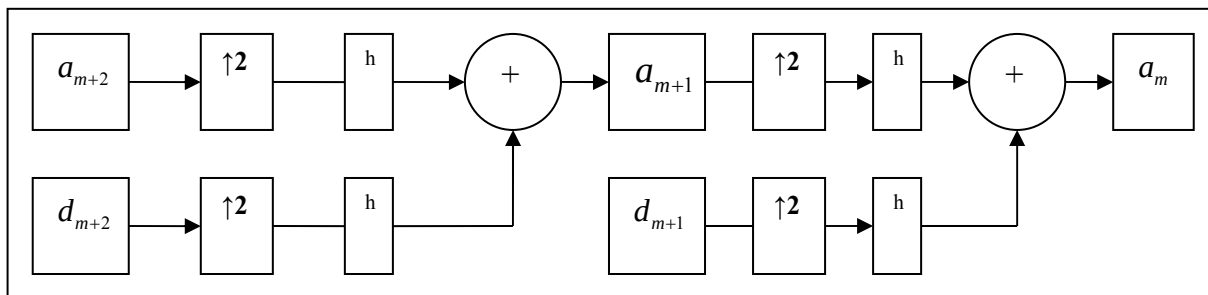


Figure 4 : Reconstitution à base d'ondelette avec filtres RIF.

Même si les ondelettes orthogonales sont un outil remarquable dans le domaine de la compression, en pratique leur implémentation pose un problème du fait de la quasi inexistence de filtres orthonormaux RIF à phase linéaire. Cela nous conduit à utiliser d'autres ondelettes présentant les mêmes avantages et pour lesquelles la construction de filtres est possible.

3. Ondelettes biorthogonales

Ce sont des ondelettes faisant intervenir pour leur construction deux familles de fonctions duales, l'une sera utilisée pour la décomposition (ψ et $\phi_{m,n}$) et l'autre pour la reconstruction ($\tilde{\psi}$ et $\tilde{\phi}_{m,n}$). Ces familles sont orthogonales entre elles mais ne le sont pas en elles-mêmes.

Elles vérifient les conditions suivantes :

1- $\psi(t)$ est orthogonal à $\tilde{\phi}_{m,n}(t)$,

2- $\phi_{m,n}(t)$ est orthogonal à $\tilde{\psi}(t)$ (d'où le nom de biorthogonalité)

Elles s'expriment dans le domaine temporel par :

$$\begin{cases} \phi(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} h_n[n] \phi(2t - n) \\ \psi(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} g_n[n] \phi(2t - n) \end{cases} \quad (3)$$

$$\begin{cases} \tilde{\phi}(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} \tilde{h}_n[n] \tilde{\phi}(2t - n) \\ \tilde{\psi}(t) = \sqrt{2} \sum_{n=-\infty}^{+\infty} \tilde{g}_n[n] \tilde{\phi}(2t - n) \end{cases} \quad (4)$$

3.1 Principe d'implémentation

Les équations (3) et (4) montrent que la décomposition et la reconstruction d'une résolution à l'autre s'implémentent par des filtres à reconstruction parfaite.

a. Mécanisme de décomposition

Des équations (3) on déduit que le calcul des coefficients d'approximation $a[m+1]$ et de détails $d[m+1]$ d'un signal à partir de ses coefficients $a[m]$ se fait par application des filtres h et g puis par un sous-échantillonnage.

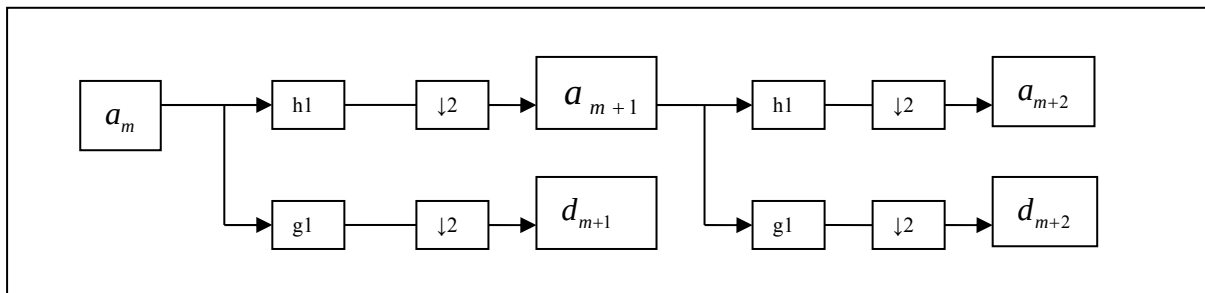


Figure 4 : Schéma de décomposition.

b. Mécanisme de reconstruction

Inversement, la reconstruction de $a[m]$ à partir de $a[m+1]$ et de $d[m+1]$ se fait en insérant un zéro entre chaque échantillon puis en faisant la somme des convolutions avec les filtres duals $\tilde{h}$ et $\tilde{g}$ associés à l'équation (4).

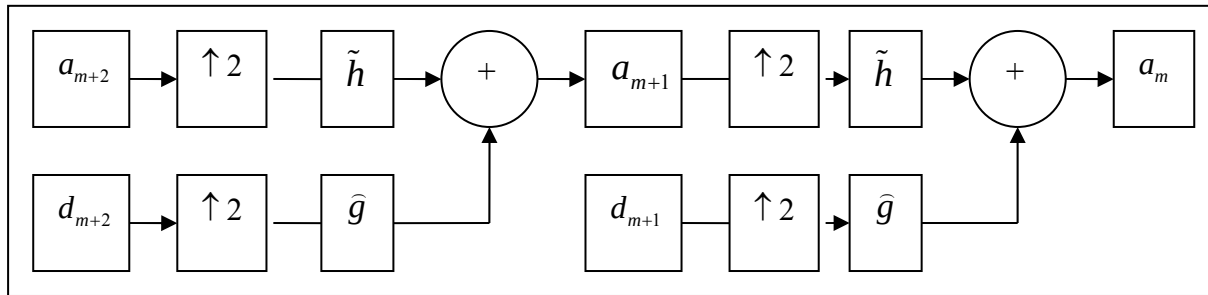


Figure 5 : Schéma de reconstitution.

L'utilisation de bases biorthogonales permet la construction de filtres associés à des ondelettes possédant un haut degré de régularité. L'analyse qui en résulte reste non redondante (comme dans le cas orthonormé, la reconstruction est assurée) et possède les propriétés de symétrie et de simplicité requises, car la résolution des filtres d'ondelettes biorthogonales se résume à la résolution d'un filtre et à la déduction de l'autre grâce aux propriétés de biorthogonalité.

Analyse en Composantes Indépendantes

L'analyse en composantes indépendantes (ICA) est une méthode récemment développée dans lequel le but est de trouver une représentation linéaire des données non gaussiennes afin que les composantes soient statistiquement indépendantes. Une telle représentation semble saisir la structure essentielle des données dans de nombreuses applications, y compris l'extraction de caractéristiques et de la séparation du signal.

1. Analyse en Composantes Indépendantes (ACI) [86].

L'ACI cherche l'indépendance des statistiques d'ordre supérieur à deux des variables mesurées. Elle est aussi connectée, en tant que méthode d'exploration de données, aux idées de la méthode dite projection-poursuite (PP). En utilisant la méthode PP, on cherche des directions intéressantes dans un nuage de données multidimensionnelles, en arguant du fait que les structures intéressantes sont celles qui font apparaître le plus l'aspect non gaussien des données projetées. Il existe un lien très significatif entre dépendance, corrélation et l'aspect non gaussien d'une variable aléatoire dans le cadre du modèle linéaire de l'ACI. En fait, la corrélation ne dépend que de la matrice de covariance. Elle représente, en quelque sorte, la partie gaussienne de l'information mutuelle qui existe entre la vraie distribution de la variable aléatoire et sa distribution estimée. L'aspect non gaussien d'une variable aléatoire est la divergence qui existe entre sa distribution et sa meilleure approximation gaussienne. Ainsi, l'ACI, qui cherche à expliquer une variable aléatoire multidimensionnelle X , en termes de composantes linéairement et statistiquement indépendantes, peut être reformulée comme étant la minimisation de la corrélation entre les composantes de X tout en maximisant l'aspect non gaussien de chaque composante.

L'analyse en composantes indépendantes d'un vecteur aléatoire réel x de taille p consiste à estimer le modèle génératif des données :

$$X = A S \dots\dots\dots (1)$$

Où: A : est la matrice de mélange et S : est un vecteur aléatoire de taille n .

Le but est donc de retrouver la matrice de séparation W telle que :

$$S = WX \dots\dots\dots (2)$$

de telle manière que les composantes de S soient aussi indépendantes que possible au sens de la maximisation d'une fonction $F(S)$ qui est une mesure de l'indépendance statistique .

Les différentes approches de l'ACI se différencient par la mesure d'indépendance retenue et par la méthode d'optimisation choisie.

2. Divergence de Kullback-Leibler

La divergence de Kullback-Leibler permet de mesurer l'écart qui existe entre la densité conjointe et le produit des densités marginales. Cet écart est la mesure d'indépendance statistique. Notons que les composantes de la variable aléatoire $X = (x_1, x_2, \dots, x_m)^T$ sont dites mutuellement indépendantes statistiquement si et seulement si :

$$P(X) = \prod_{i=1}^m P(x_i) \text{ pour chaque paire } i, j \text{ } 1 \dots m \dots\dots\dots (3)$$

où $P(X)$ est la distribution de X et $\prod_{i=1}^m P(x_i)$ est le produit de ses distributions marginales.

Soit $f(x)$ et $g(x)$ deux densités de probabilités des deux composantes de la variable aléatoire X . Ainsi, la divergence de Kullback-Leibler est définie par :

$$D_{f,g} = \int_{-\infty}^{+\infty} f(x) \log\left(\frac{f(x)}{g(x)}\right) dx \dots\dots\dots (4)$$

Il est clair que la divergence de Kullback-Leibler ne s'annule que lorsque les deux composantes de X sont statistiquement indépendantes.

3. Information mutuelle

L'information mutuelle peut être considérée comme la distance de Kullback-Leibler entre les densités conjointes et les densités marginales. Ainsi, l'information mutuelle de la variable aléatoire X , notée $I(X)$, peut être exprimée comme étant la divergence de Kullback-Leibler entre sa distribution $P(X)$ et le produit de ses distributions marginales $P(x_i)$ tel que :

$$I(x) = D(P(X) \prod_{i=1}^m P(x_i)) \dots\dots\dots (5)$$

En utilisant le concept de l'entropie différentielle, on peut reformuler l'information mutuelle $I(X)$ comme suit :

$$I(X) = I(x_1, \dots, x_n) = \sum_{i=1}^n H(x_i) - H(X) \dots\dots\dots (6)$$

où les $H(x_i)$ sont les entropies marginales des composantes x_i et $H(X)$ l'entropie conjointe.

4. Principaux algorithmes

Le but de cette section est de donner les principales idées des algorithmes les plus connus permettant de résoudre le problème d'ACI. Ces algorithmes ont été choisis de manière à présenter un panel varié d'approches. On trouvera de nombreuses autres méthodes dans la littérature scientifique.

a) Algorithme d'Herault et Jutten :

Le premier algorithme d'ACI (du nom de ses auteurs) procède par une approche neuromimétique: il s'agit d'une utilisation des outils du traitement du signal pour modéliser un phénomène inspiré du fonctionnement neuronal. Cet algorithme était de type Robins-Monro, et recherchait itérativement des zéros communs de deux fonctions. Il est basé sur un réseau de neurones récurrent dont les poids sont les termes non diagonaux de la matrice de séparation W , les termes diagonaux étant contraints à la nullité. L'estimation des sources est ainsi donnée par:

$$y = (I+W)^{-1}x \dots\dots\dots (7)$$

Avec la règle d'adaptation suivante pour les termes non diagonaux:

$$\Delta w_{i,j} = f(y_i) \cdot g(y_j) \dots\dots\dots (\text{III. } 8)$$

f et g étant des fonctions non linéaires impaires différentes à choisir en fonction des densités de probabilité des sources à estimer. Des travaux ultérieurs ont justifié la forme de cette règle d'apprentissage et ont montré que la fonction non linéaire devait être choisie au sens du maximum de vraisemblance. La mesure d'indépendance sous-jacente à cet algorithme est l'annulation des cumulants d'ordre supérieur.

b) Maximisation de Contraste (CoM)

Cette famille d'algorithmes a été proposée par Comon en 1991. Ces algorithmes procèdent en balayant toutes les paires de sorties à tour de rôle. L'algorithme CoM peut calculer soit tous les cumulants des observations (comme JADE), soit à la demande, à chaque fois qu'une rotation optimale est calculée. Le choix peut être fait de manière à minimiser la complexité numérique. Plusieurs versions ont été proposées, suivant le contraste utilisé. En 1991, il s'agissait par exemple de la somme des modules carrés des cumulants marginaux des sorties: ce contraste est maintenant appelé CoM2. Plus tard, un autre contraste, CoM1, a été proposé par Moreau: la

somme des modules des cumulants marginaux des sorties. Lorsque les sources ont des cumulants marginaux toutes de même signe, il a été montré que la valeur absolue peut être retirée. Cette simplification a donné lieu à une autre solution algébrique compacte, développée par Comon.

D'autres formes plus générales de contraste basées sur les cumulants ont été ensuite proposées dans la littérature .

c) JADE (Joint Approximate Diagonalization of Eigenmatrices):

Un tenseur de cumulants (à l'ordre quatre) est une matrice à quatre dimension contenant tous les cumulants croisés d'ordre quatre. C'est aussi une application linéaire d'un espace de matrice de taille $N \times N$ dans un autre espace de matrice de même taille. La diagonalisation de cette application linéaire permet d'effectuer la séparation voulue sous certaines conditions. En effet, l'indépendance statistique est vue ici comme la recherche à tendre vers la nullité de tous les moments et cumulants (à tous les ordres), ce qui revient à annuler tous les éléments non diagonaux de la matrice associée à l'application linéaire sus-citée. Cardoso et Souloumiac ont proposé une méthode de diagonalisation jointe permettant de mettre en pratique ce principe. Cela revient à minimiser la somme du carrés de tous les cumulants (à l'ordre quatre) « hors diagonale » (cumulant entre deux signaux différents). En pratique, JADE nécessite le calcul de tous les cumulants d'ordre quatre et a donc une complexité en $O(n^4)$.

5. Prétraitement

Toutefois, avant d'appliquer un algorithme d'ACI sur les données, il est généralement très utile de faire un prétraitement afin de simplifier la théorie et les algorithmes d'ACI. Dans cette section, nous discutons de certaines techniques de prétraitement qui rendent le problème de l'estimation des composantes indépendantes plus simple et mieux conditionnés.

a) Centrage

Le prétraitement le plus fondamental et nécessaire est de forcer les composantes du vecteur d'observations x à une moyenne nulle. Ceci est lié aux cumulants croisés et à leurs estimateurs qui sont d'expression beaucoup plus simple dans le cas de variables aléatoires centrées. Cette étape est réalisée par une simple soustraction de la moyenne des différentes composantes à analyser. La matrice de mélange A n'est pas modifiée par cette opération qui peut donc toujours être appliquée aux données sans affecter l'estimation de A .

b) Blanchiment des données :

Une autre stratégie très utile pour le pré-traitement à l'ICA est le blanchiment des données, cette étape consiste à décorrélérer et imposer une variance unité aux variables du vecteur d'entrée. Cette transformation linéaire est réalisée en multipliant le vecteur d'entrée centré x^c par

une matrice V de dimension $P \times N$ afin d'obtenir le vecteur : $Z = Vx^c$ de dimension $P \times 1$ où les composantes de Z sont décorréliées et de variance unité, ce qui signifie que la matrice de covariance de Z vaut :

$$R_Z = E\{Z Z^T\} = I_p \dots \dots \dots (III. 9)$$

Où I_p est la matrice identité de dimension $P \times P$.

Diverses méthodes ont été développées pour réaliser cette transformation et l'une des plus populaires est l'utilisation de l'Analyse en Composantes Principales (ACP).

L'ACP d'un vecteur aléatoire réel x de taille p et de matrice de covariance $V_x = E\{x.x^T\}$ est définie comme un couple de matrice $\{F, D\}$ tel que la matrice de variance/covariance se factorise sous la forme :

$$V_x = F.D.F^T \dots \dots \dots (III. 10)$$

D est une matrice diagonale réelle positive

F est une matrice de rang r et de taille $p \times r$ dont les colonnes sont orthogonales entre elles (c'est-à-dire que $F^T.F$ est une matrice diagonale).

Une méthode pratique pour réaliser une ACP est donc de diagonaliser la matrice de covariance des données et de définir la matrice D comme une matrice diagonale contenant les valeurs propres non nulles de V_x rangées dans l'ordre décroissant et F telles que ses colonnes contiennent les vecteurs propres correspondants.

Dans le cas d'une diagonalisation de la matrice de covariance, les vecteurs propres ont une norme unitaire, si bien que $F^T.F$ est égale à la matrice unité. Ainsi la projection des données sur le premier vecteur propre, appelée première composante principale, encode un maximum de variance. Les composantes principales suivantes sont déterminées de telle façon qu'elles encodent le maximum de la variance restante. Dans le cas de réduction de dimension, seules les p premières valeurs propres et les vecteurs associés sont conservés.

En définissant la matrice diagonale $D^{1/2}$ de dimensions $p \times p$ $diag(d_1^{-1/2}, \dots, d_n^{-1/2})$.

La matrice de blanchiment B de dimensions $p \times n$ est calculée par :

$$B = D^{1/2} \times E^T \dots \dots \dots (11)$$

Le blanchiment des données n'est pas systématique. Certains algorithmes d'ACI s'appliquent directement sur les données centrées. Il est cependant recommandé puisqu'il permet d'accélérer la convergence des algorithmes .

Algorithme de blanchiment

- Calcul de la matrice de covariance des données : $cov(X) = X^T X$

- Diagonalisation de la matrice de covariance : $\text{cov}(X) = D * E$ où D est une matrice diagonale composée des valeurs propres et E est composée des vecteurs propres de la matrice de covariance
- Calcul de « l » le nombre de composantes à garder.
- Réduction de la matrice caractéristique = $[V_1, V_2, \dots, V_l]$
- Blanchiment des données : $Xb = B * X$

6. Eloignement à la gaussianité :

L'Analyse en Composante Indépendantes peut être abordée, par la recherche de distributions les plus éloignées possibles de la distribution normale. La justification essentielle de ce point de vue est le théorème central limite qui stipule que la somme de variables indépendantes tend asymptotiquement vers une distribution normale. De plus, la somme de deux variables aléatoires indépendantes est généralement plus proche de la gaussienne que n'importe laquelle des deux variables. Les observations qui sont des combinaisons linéaires des sources sont alors davantage gaussiennes que ces dernières, et l'estimation de W^{-1} peut être basée sur la maximisation de la non gaussianité des sources. Une mesure de cette non gaussianité s'impose donc. La plus classique est le Kurtosis qui est le cumulants normalisés d'ordre 4, mais il reste peu robuste:

$$Kurt(y) = E\{y^4\} - 3E\{y^2\}^2 \dots\dots\dots (12)$$

La seconde mesure trouve ses bases dans la théorie de l'information. Il est établi que pour un ensemble de variables aléatoires de même variance, celle qui suit la loi gaussienne maximise l'entropie donnée par :

$$H(y) = - \int f_y(u) \log f_y(u) du \dots\dots\dots (13)$$

où $f_y(y)$ est la densité de y .

Cette remarque permet d'introduire une mesure de la propriété non gaussienne appelée néguentropie définie par :

$$J(y) = H(y_g) - H(y) \dots\dots\dots (14)$$

où y_g est un vecteur aléatoire gaussien de même covariance que y .

La néguentropie est égale à la divergence de Kullback-Leibler entre $f_y(y)$ et $f_{y_g}(y_g)$. Etant donnée la difficulté d'évaluer J d'après l'équation (III.14), cette quantité est en pratique approximée. Hyvärinen a proposé une approximation robuste de la néguentropie :

$$J(y) = [E\{G(y)\} - E\{G(y_g)\}]^2 \dots\dots\dots (15)$$

où G est une fonction non quadratique.

Hyvärinen a proposé un algorithme robuste, baptisé fastICA, qui maximise un contraste basé sur la néguentropie pour l'optimisation de la matrice de séparation.

7. Fast ICA

FastICA, est un algorithme très performant de maximisation de la fonction de contraste pour les sources nongaussiennes. Elle est basée sur l'algorithme d'apprentissage itératif de point fixe (*fixed point algorithm*).

Comme l'indépendance implique la décorrélation, la maximisation de la néguentropie se fait sur l'espace réduit des données décorrélatées par une PCA initiale où toutes les valeurs propres sont gardées. L'algorithme va suivre les étapes suivantes après un blanchiment des données :

- 1- Choisir un vecteur initial w aléatoire.
- 2- Mettre $w^+ = E(x.g(w^T x)) - w.E(g_0(w^T x))$.
- 3- Mettre $w = w^+ / \|w^+\|$.
- 4- Si non convergence, aller à 2.

La convergence signifie que l'ancienne et la nouvelle valeur de w vont dans la même direction. Comme pour beaucoup d'algorithmes, les données sont contraintes à être centrées et décorrélatées. Dans le cas où l'on considère toutes les colonnes à la fois, il faut ajouter une contrainte de blanchiment global à chaque itération. Le point clé est que sous contraintes de blanchiment, la décorrélation est équivalente à une orthogonalisation. Cette orthogonalisation évite que les différentes colonnes convergent vers la même source. Deux stratégies peuvent alors être choisies pour contraindre à l'orthogonalité. La première se base sur la méthode de Gram-Schmidt ; dans cette méthode, après l'estimation de p composantes indépendantes $(w_1, \dots, w_p)$, on utilise l'algorithme de point fixe pour estimer la $(p+1)^{\text{ème}}$ composante, i.e. vecteur w_{p+1} , et après chaque itération on soustrait du vecteur w_{p+1} les projections $w_{p+1}^T w_j w_j$, ($j=1, \dots, p$), après on normalise w_{p+1} .

-Mettre $w_{p+1} = w_{p+1} - \sum_{j=1}^p w_{p+1}^T w_j w_j$

-Mettre $w_{p+1} = w_{p+1} / \sqrt{w_{p+1}^T w_{p+1}}$

L'inconvénient de cette méthode, appelée « approche par déflation », est qu'une erreur d'estimation sur une composante se répercute sur les suivantes du fait de cette orthogonalisation. L'autre méthode, appelée « approche symétrique », orthogonalise simultanément toutes les colonnes en utilisant une décorrélation symétrique de type $W=(WW^T)^{1/2}.W.$, où W est la matrice représentant les $w_1, \dots, x_n$ vecteurs. La racine carrée inverse $(WW^T)^{1/2}$ est obtenue de la décomposition en valeurs propres de $WW^T=F.A.F^T$ où A est la matrice diagonale des valeurs propres et F la matrice des vecteurs propres correspondants. On a alors :

$$(WW^T)^{1/2}=F.A^{-1/2}.F^T \dots\dots\dots (15)$$

Elle réclame donc plus de mémoire, mais n'est pas pour autant moins rapide

a) L'analyse en Composantes Indépendantes par l'approche symétrique :

Initialisation de la matrice appliquant l'ACI : $W^+ \longrightarrow W_0^+$

Tant que la condition de fin n'est pas remplie :

Orthogonalisation de la matrice appliquant l'ACI : $(W_n^+)^{\perp} = W_n^+((W_n^+)^T W_n^+)^{-1/2}$

si $\|W_n^+ - W_{n-1}^+\| \leq \varepsilon$

La condition de fin est remplie

Sinon

si le numéro de l'itération est égal au nombre maximal d'itération

La condition de fin est remplie

Sinon

Le numéro de l'itération est incrémenté

Mise à jour de la matrice $W^+ : W_{n+1}^+ = \frac{1}{N} . Xb([y_{i,j}^3]_n) - 3W_n^+$

Application de l'ACI : $Xaci = W^+ Xb$