

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
UNIVERSITE AKLI MOAND OULHADJ-BOUIRA



Faculté des Sciences et Sciences Appliquées
Département Génie Electrique

Mémoire de fin d'étude

Présenté par :

AKKAL Sabrina

CHEDDANI Karima

En vue de l'obtention du diplôme de **Master en :**

Filière : Génie biomédical

Option : Instrumentation biomédical

Thème :

Analyse de l'électrocardiogramme humain pour la classification automatique des arythmies et les systèmes de reconnaissance biométrique à l'aide de paramètres de modélisation analytiques et autorégressifs.

Devant le jury composé de :

Mr. BENZAOUI	MCA	UAMOB	Président
Mr. BENGHANIA	MAA	UAMOB	Encadreur
Mr. DIB	MAA	UAMOB	Examinateur

Année Universitaire 2019/2020



Remerciements

*Avant tout, je remercie Dieu de m'avoir donné la force pour
réaliser ce présent travail.*

(الحمد لله)

*Je remercie tout particulièrement "Mr benghenia hadj abd el
kader" de m'avoir encadré durant ces années de master en étant
toujours disponible et encourageant, pour son aide et conseils et
pour ses grandes valeurs humaines.*

*Nos respectueux remerciements à l'ensemble des membres du
jury pour l'intérêt porté sur notre travail.*

*Merci à tous ceux qui m'ont aidé de près ou de loin durant ces
années de recherche et d'études. Je ne peux conclure sans un mot
de remerciement à tous mes enseignants du Département de
Génie Electrique, Université de Bouira tout le long de ma
formation.*

Je vous remercie !



Dédicace

Je rends grâce a Dieu de m'avoir donné le courage et la volonté ainsi que la conscience d'avoir pu terminer mes études.

Je dédie ce modeste travail :

A ma mère adorée qui a su se montrer patiente et qui n'a jamais cessé de ménager leurs efforts pour que j'atteigne ce niveau.

A l'esprit de mon père (rahimaho Allah)

Qu'ils trouvent en ce travail la preuve de tout mon amour et ma tendresse

A mon marie RAMDHAN a été toujours le pour me donner les conseils et pour m'encourager le long de ces années avec tout son amour et a toute sa famille

A mes chers sœurs :LUIZA, HANANE,NAIMA et les petits MARIA ,AYOUB, ISLAM ,JYAD,SAMI AMIR

A tout mes amis(es) surtout SABRINA

Tous ce que j'aime et qui m'aiment

MERCI

KARIMA

Dédicace

Avec un grand cœur plein de chaleur, je dédie ce mémoire, aux êtres qui me sont les plus chers:

A mon cher père

Pour son soutien inconditionnel, ses encouragements et pour m'avoir permis de réaliser mes études dans les meilleurs conditions.

A ma chère mère

Pour m'avoir soutenu, accompagné et surtout encouragé tout au long de ce travail.

A mes sœurs Iman et Lamia et les petits Adame et maria.

A mon frère Mohammed.

Et à toute ma Famille.

A mes amis surtout Karima.

A toutes les personnes que j'ai rencontrées durant toute la vie.

A toute la promo instrumentation biomédicale 2020.

A tous ceux qui m'ont soutenu pendant toute la durée de mes études.

SABRINA

Sommaire

Remerciement	
Dédicace	
Résumé	
Sommaire.....	I
Liste des figures.....	V
Liste des Tableaux.....	VIII
Liste des significations	IX
Introduction générale	01

Chapitre 1: Notions de base de l'électrocardiographie

1.1	Introduction.....	03
1.2	Historique de l'électrocardiographie.....	03
1.3	Système cardiovasculaire	04
1.3.1	Anatomie du cœur humain.....	04
1.3.2	La circulation sanguine dans le cœur.....	05
1.3.3	L'activité électrique du cœur.....	06
1.4	Le signal électrocardiogramme (ECG).....	08
1.4.1	les dérivations électrocardiographiques.....	09
1.4.1.1	les dérivations périphériques.....	10
1.4.1.2	les dérivations précordiales.....	11
1.5	Les arythmies cardiaques.....	12
1.5.1	La bradycardie.....	12
1.5.2	La tachycardie	13
1.5.3	L'extrasystole.....	13
1.5.4	La fibrillation auriculaire	14
1.5.5	La fibrillation ventriculaire	14
1.5.6	Bloc de branche.....	15
1.5.6.1	Bloc de branche droite.....	15
1.5.6.2	Bloc de branche gauche.....	16
1.6	Artefacts visibles sur l'électrocardiogramme.....	17
1.6.1.	Bruits techniques	17
1.6.2	Artefacts physiques	19

Sommaire

1.7	La base de donnée MIT/BIH	21
	Conclusion.....	22
Chapitre2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)		
2.1	Introduction.....	23.
2.2	Méthodes de filtrage de signal ECG.....	23
2.2.1	Filtrage de la dérive de la ligne de base.....	23
2.2.2	Filtrage du bruit d'interférence à 50 Hz (ou 60 Hz)	24
2.2.3	Filtrage du bruit de l'électromyogramme	24
2.3	Détection les ondes P, R, Q, S, et T	25
2.3.1	Filtrage de signal ECG	25
2.3.2	Détection des pics R.....	25
2.3.3	Détection des ondes Q et S.....	26
2.3.4	Détection des ondes T.....	26
2.3.5	Détection des ondes P.....	27
2.4	Les différents types des modèles autorégressifs et linéaires.....	27
2.5	Algorithmes d'identification paramétrique.....	28
2.5.1	Sélection de la classe du modèle.....	29
2.5.2	Sélection de l'ordre.....	30
2.5.3	Procédure de validation du modèle.....	31
2.5.4	Extraction des paramètres du modèle.....	31
2.6	Modèles paramétriques autorégressifs et linéaires.....	32
2.6.1	Modèle Autorégressif (AR).....	33

Sommaire

2.6.2 Modèle à moyenne ajusté (Moving Average MA).....	34
2.6.3 Modèle ARMA.....	34
Conclusion	35

Chapitre 3 : Méthodes des fonctionnalités d'extraction pour les signaux ECG

3.1 Introduction.....	36
3.2 Méthode d'extraction du système biométrique ECG.....	36
3.3 Caractéristiques analytiques.....	38
3.4 Fonctionnalités de modélisation.....	42
Conclusion	45

Chapitre 4 : Classification SVM

4.1. Introduction.....	46
4.2. Notions sur l'apprentissage statistique.....	46
4.3. Théorie des machines à vecteurs de support (SVM)	46
4.4. Principe des machines à vecteurs de support	47
4.4.1. Cas de données linéairement séparable.....	47
4.4.2. Cas des données non-linéairement séparables	47
4.5. Classification SVM	48
4.5.1. Formulation d'un problème de classification (supervisé)	49
4.5.2. Minimisation du risque structurel :	49
4.5.3. Architecture d'un classificateur SVM :	50
4.5.3.1. La fonction noyau	50
4.5.3.2. Sélection de modèle SVM :	51
4.5.3.3. Estimation de l'erreur de généralisation :	51
4.6. Les Machines à Vecteurs Supports (SVM) pour la régression	52
4.6.1. Formulation du problème dans le cas linéaire :.....	53
4.6.2. Extension au cas non linéaire:	53
4.6.3. Choix des hyperparamètres:	53
4.7. Résultats.....	54

Sommaire

Conclusion.....55

Conclusion général.....56

Résumé

Liste des figures

Chapitre 1: Notions de base de l'électrocardiographie

Figure 1.1 Schéma fonctionnel du cœur.....	04
Figure 1.2 La circulation sanguine.....	06
Figure 1.3 Activité électrique du cœur.....	07
Figure 1.4 Le signal ECG.....	08
Figure 1.5 Montage d'Einthoven pour l'enregistrement des dérivations bipolaire Des membres.....	10
Figure 1.6 Montage d'Einthoven pour l'enregistrement des dérivations unipolaires Des membres.....	11
Figure 1.7 Dérivations thoraciques.....	12
Figure 1.8 Bradycardie sinusale.....	12
Figure 1.9 Rythme d'échappement jonctionnel à 35 bpm.....	13
Figure 1.10 Tachycardie sinusale.....	13
Figure 1.11 Fibrillation auriculaire.....	14
Figure 1.12 Fibrillation ventriculaire.....	15
Figure 1.13 Bloc de branche droit représenté par les 12dérivtions.....	16
Figure 1.14 Bloc de branche gauche représenté par les 12 dérivations.....	16
Figure 1.15 Interférence secteur 50 Hz.....	19
Figure 1.16 Bruit dû aux mouvements des électrodes.....	19
Figure 1.17 Dérive de la ligne de base.....	20
Figure 1.18 Bruit myoélectrique ou tremblement somatique.....	21

Liste des figures

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

Figure 2.1 La procédure générale d'identification	29
Figure 2.2 Optimisation de l'ordre de prédiction (a) FPE, (b), AIK, (c) évolution De l'erreur.....	31
Figure 2.3 Structure des modèles paramétriques autorégressifs et linéaires.....	32
Figure 2.4 Structure du modèle AR.....	34
Figure 2.5 Structure du modèle MA.....	34
Figure 2.7 Structure du modèle ARMAX.....	35

Chapitre 3 : Méthodes des fonctionnalités d'extraction pour les signaux ECG

Figure 3.1 Schéma fonctionnel du système d'identification proposé basé Sur l'ECG.....	36
Figure 3.2 (a) signal ECG bruité, (b) signal ECG filtré (enregistrement 101).....	38
Figure 3.3 détection des pics R.....	39
Figure 3.4 détection des ondes P.....	39
Figure 3.5 détection des ondes Q.....	40
Figure 3.6 détection des ondes Q.....	40
Figure 3.7 détection des ondes T.....	41
Figure 3.8 La variabilité des intervalles cardiaque par rapport RRI.....	41
Figure 3.9 Le signal RR original (trait continu), et signal RR simule (Trait discret) en utilisant le modèle AR.....	43
Figure 3.10 La DSP du signal RR original (trait continu) et signal RR simule (Trait discret) en utilisant le modèle AR.....	43

Liste des figures

Figure 3.11 Espace des caractéristiques 3D des coefficients AR extraits du Signal RRI résultant en une classification complètement précise.....44

Chapitre 4 : Classification SVM

Figure 4.1 Cas linéairement séparable.....47

Figure 4.2 Cas non linéairement séparable.....48

Figure 4.3 Architecture d'une machine à vecteurs de support (d'un nombre N)...50

Figure 4.4 Coefficients A1 et A2 AR extraits résultant en une classification complètement précise à l'aide de SVM.....54

Figure 4.5 Coefficients A1 et A3 AR extraits résultant en une classification complètement précise à l'aide de SVM.....54

Liste des tableaux

Tableau	Titre	page
Tableau 2.1	Différents modèles autorégressifs et leurs coefficients.....	33
Tableau 3.1	Liste des fonctionnalités extraites.....	37
Tableau4.1	La précision (%) obtenu dans la classification correspondant à l'algorithme biométrique.....	55

Liste des abréviations

A1 : Le premier coefficient de modèle AR

A2 : Le deuxième coefficient de modèle AR

A3 : Le troisième coefficient de modèle AR

AIC : Critère d'information d'Akaike

AR : Le modèle autorégressif

ARMA : Autoregressive Moving Average

AV: Auriculo-ventriculaire

Bpm : Battement par minute

DI, DII, DIII : Dérivations bipolaires

DSP : power spectral density (densité spectrale de puissance)

ECG : Electrocardiogramme

EPF : Critère de l'erreur de prédiction finale

FV : Fibrillation ventriculaire

GLS : Les moindres carrés généralisés

LBBB : Bloc de branche gauche

LOO : Leave one out

LS : Critère des moindres carrés

MA : Modèle à moyenne ajusté (moving average)

MIMO : Multiple input/ Multiple output

MIT : Massachusetts institute of technology

P : L'onde P de signal ECG

PVC : Contraction ventriculaire prématurée

QRS : Trois ondes Q, R, S

Liste des abréviations

RBF : Radial Basis Function

SISO : Single input/Single output

SVM : Machine à vecteurs de supports

SVR : Support Vector Regression

T : L'onde T de signal ECG

VCI: Veine cave inférieure

VCS : Veine cave supérieure

VF : Potentiel sur la jambe gauche

VL : Potentiel sur le bras gauche

VR : potentiel sur le bras droit

Introduction générale

Introduction générale

La première cause de mortalité dans le monde, provient des maladies cardiovasculaires. Même si les connaissances acquises en cardiologie sont grandes le cœur n'a pas dévoilé tous ses secrets. Pour cette raison, il est important de traiter et d'analyser le signal électrique décrivant l'activité cardiaque. Un tel signal : le signal Electrocardiogramme, englobe différents paramètres permettant de décrire différents problèmes cardio-vasculaires tout en intégrant le côté physiologique.

Le signal électrocardiogramme (ECG) est l'outil de diagnostic utilisé pour évaluer la probabilité d'anomalies cardiaques. Il enregistre les impulsions électriques qui déclenchent la contraction cardiaque et donne des informations sur la phase de récupération. Il est constitué d'un ensemble d'ondes, P, Q, R, S, T, et un ensemble d'intervalles, l'intervalle RR, l'intervalle QT, l'intervalle PR, le segment ST, et le complexe QRS. Toute l'information utile d'un signal ECG est contenue dans ses points extrêmes ou pics, et le changement des intervalles indiquant si ce dernier est normal ou présentant des anomalies [1].

Les signaux ECG dans ce travail sont étudiés afin de concevoir deux applications basées sur l'ECG, de classer et de reconnaître les signaux ECG normaux (sains) des signaux ECG anormaux (malsains) et d'identifier chaque personne individuellement en concevant un système biométrique basé sur l'ECG. Le traitement automatique du signal et l'analyse des signaux ECG ont été introduits en tant qu'outil pour les applications de détection d'arythmie et de surveillance de signal. Ces dernières années, un certain nombre de techniques de traitement du signal pour l'auto-classification du signal ECG et les systèmes d'identification biométrique ont été proposés et évalués.

Le but des techniques dans le contexte du traitement et de la classification ECG est d'obtenir la précision de classification la plus élevée avec le plus petit nombre de paramètres d'extraction du signal ECG afin d'améliorer la productivité du système. Dans ce travail, des techniques de modélisation sont appliquées pour extraire des caractéristiques ou des paramètres de battements ECG uniques ou multiples afin d'utiliser ces paramètres pour deux tâches de classification - la classification ECG et la reconnaissance biométrique ECG. Les modèles autorégressifs sont utilisés comme technique d'extraction à des fins de classification ECG.

Introduction générale

Afin de garantir un taux de modélisation de haute précision pour le signal ECG, il est fondamental d'analyser le signal ECG en utilisant l'ordre qui peut produire une précision de modélisation élevée du signal ECG pour chaque processus de modélisation. Sans modélisation précise du signal ECG, les caractéristiques extraites du modèle appliqué ne s'attendent pas à montrer des résultats de classification élevés. Dans ce travail, la sélection de l'ordre du modèle appliqué peut être estimée à l'aide de techniques qui comparent les erreurs de modélisation de différents ordres afin d'estimer l'ordre avec l'erreur de modélisation minimale.

La classification ECG proposée et les systèmes biométriques sont mis en œuvre en utilisant AR (3), c'est-à-dire $p = 3$. Les paramètres autorégressifs sont ensuite extraits de chaque battement du flux ECG et intégrés dans des algorithmes de classification pour évaluer le taux de classification entre les signaux ECG normaux et anormaux ou pour identifier chaque signal ECG individuellement. La méthode proposée d'utilisation de modèle autorégressif pour concevoir un système de classification ECG ont montré des résultats de classification très élevés en termes de précision. La classification ECG et le système biométrique conçus se composent de quatre étapes principales, illustrées à la **Figure 3.1**.

En résumé, le travail de ce mémoire est réparti en quatre chapitres :

Après l'introduction générale,

- Le premier chapitre, va décrire les généralités sur l'électrophysiologie cardiaque.
- Le deuxième chapitre, va être consacré à la présentation des différents algorithmes utilisés pour la détection des différentes ondes et intervalles de signal ECG. Nous avons aussi présenté les trois différents modèles que nous avons utilisés tel que le modèle AR, MA, et ARMA.
- Le troisième chapitre, s'intéresse au traitement du signal ECG. Il est réparti en deux parties : la première partie traite le filtrage, la et la détection du complexe QRS, l'onde P et l'onde T. La deuxième partie traite l'extraction des paramètres du modèle estimé (AR modèle). Des résultats sont présentés au niveau de chaque partie.
- Dans le quatrième chapitre, la classification est réalisée en utilisant la technique machines à vecteurs de support (SVM), et nous allons évaluer notre algorithme selon des critères théorique qui mesure le taux de classification, la sensibilité et spécificité.

Le mémoire se termine par une conclusion générale.

Chapitre 1: Les notions de base de l'électrocardiographie

1.1 Introduction

L'électrocardiographie est un examen fondamental et irremplaçable en cardiologie .Il permet d'obtenir un tracé qui traduit l'activité électrique du cœur. Dans ce chapitre des notions de base sur le système cardiovasculaire, et sur le signal électrocardiogramme ECG, les différentes arythmies cardiaques pour une meilleure compréhension du lecteur de la suite de notre travail, les différents bruits qui peuvent atteindre le signal ECG.

1.2. Historique de l'électrocardiographie

La période 1842 - 1942 constitue le premier centenaire de l'électrocardiographie et des premières découvertes scientifiques de l'ECG à 12 dérivations :

Carlo Matteucci (1842), physicien italien, montre qu'un courant électrique passe à chaque battement cardiaque.

Rudolph von Koellicker & Henrich Mueller (1856) enregistrent le premier potentiel d'action, enregistrement fait en mettant en contact un nerf de patte de grenouille avec un cœur isolé : à chaque battement cardiaque, on observait une contraction de la patte de grenouille.

Augustus D. Waller (1887), physiologiste anglais, publie le premier électrocardiogramme. En se liant à un tube capillaire, dans un champ magnétique une électrode posée sur le thorax, on voyait varier le niveau liquide dans le tube à chaque battement de cœur.

Willem Einthoven (1893) utilise le terme "électrocardiogramme" pour la première fois à une réunion de la "Deutsch Medical Association".

W. Einthoven (1895) met en évidence 5 ondes sur le tracé électrocardiographique, ondes qu'il nomme respectivement P, Q, R, S, T.

La position des dérivations précordiales V1 à V6 est définie par convention par La "American Heart Association" et la "Cardiac Society of Great Britain" en 1938.

E. Goldberger (1942) ajoute 3 dérivations périphériques aux 3 dérivations déjà décrites par Einthoven, à savoir aVR, aVL et aVF [2].

1.3 Système cardiovasculaire

Le système cardiovasculaire est constitué du cœur et du système vasculaire. Sa fonction principale est d'assurer la circulation du sang dans l'organisme afin de satisfaire aux besoins énergétiques et au renouvellement cellulaire, quelles que soient les conditions ambiantes et l'activité de l'individu [3].

1.3.1 Anatomie du cœur humain

Le cœur est un organe creux et musculaire comparable à une pompe, qui assure la circulation du sang dans les veines et les artères. Dans le corps humain, le cœur se situe un peu à gauche du centre du thorax, en arrière du sternum. Il est l'élément central du système cardiovasculaire. Il est connecté au reste de l'organisme par le biais de vaisseaux associés : les deux veines caves (inférieure et supérieure), les artères pulmonaires, et les artères aorte, comme l'illustre (la figure 1.1).

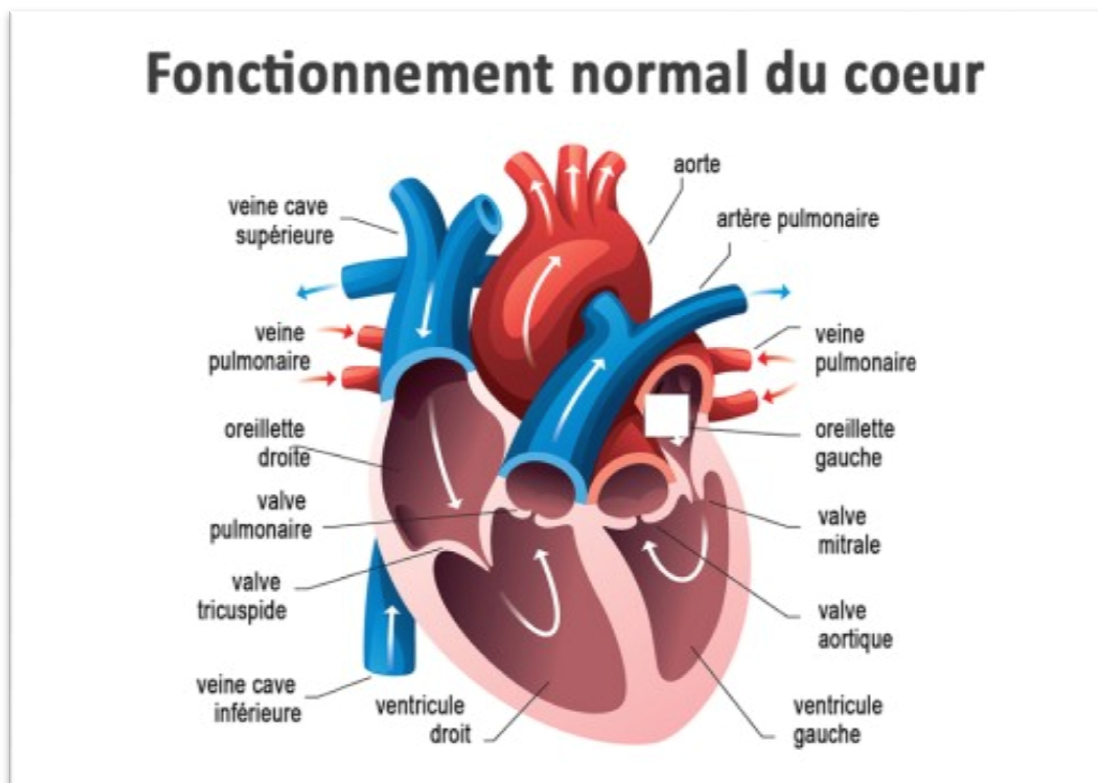


Figure 1.1. Schéma fonctionnel du cœur [3].

Le cœur se décompose en quatre chambres : les oreillettes en haut, et les ventricules en bas. Une cloison musculaire épaisse, le septum, divise le cœur en deux morceaux. Chaque côté, droit et gauche, est composé de deux cavités : l'oreillette et le ventricule, reliés entre eux

Chapitre 1: Les notions de base de l'électrocardiographie

par une valve qui assure, à l'état normal, un passage unidirectionnel du sang. Pour la cavité de droite, cette valve d'admission est la valve tricuspide ; pour la cavité de gauche c'est la valve mitrale.

Il existe aussi des valves d'échappement qui assurent la communication entre le ventricule droit et l'artère pulmonaire (valve pulmonaire), ainsi qu'entre le ventricule gauche et l'artère aorte (valve aortique). A l'état normal, ces valves d'admission et d'échappement empêchent le sang de refluer et ne peuvent être ouvertes en même temps.

Le cœur est donc séparé en deux moitiés indépendantes (droite et gauche), chacune composée d'une oreillette et d'un ventricule. Sa partie droite contient du sang pauvre en oxygène et assure la circulation pulmonaire; sa partie gauche renferme du sang riche en oxygène, et le propulse vers le reste du corps, hormis les poumons.

Les parois du cœur sont constituées par un tissu musculaire, appelé le myocarde, qui est composé de cellules musculaires cardiaques spécialisées, les cardiomyocytes, qui ne ressemblent à aucun autre tissu musculaire du corps. En particulier, ces cellules sont intétanisables, ce qui signifie qu'elles sont incapables de contraction prolongée. De plus ces cellules sont excitables, automatiques et indépendantes, conductrices et contractiles.

Les tissus annexes de myocarde sont l'endocarde, qui la tapisse à l'intérieur, et le péricarde, qui l'entoure à l'extérieur.

Les ventricules ont pour fonction de pomper le sang vers le corps ou vers les poumons. Leurs parois sont alors plus épaisses que celles des oreillettes, et la contraction des ventricules est plus importante pour la distribution du sang. Le ventricule gauche est bien plus massif que le droit parce qu'il doit exercer une force considérable pour forcer le sang à travers tout le corps contre la pression corporelle, tandis que le ventricule droit ne dessert que les poumons [3].

1.3.2 La circulation sanguine dans le cœur

Le sang veineux arrive de la périphérie par la veine cave inférieure (VCI) et veine cave supérieure (VCS), pour se jeter dans l'oreillette puis le ventricule droit.

De là il emprunte l'orifice pulmonaire puis l'artère pulmonaire pour rejoindre le lit capillaire pulmonaire ou se feront les échanges gazeux.

Chapitre 1: Les notions de base de l'électrocardiographie

Après oxygénation, le sang rejoint l'oreillette gauche par l'intermédiaire de 4 veines pulmonaires, puis le ventricule gauche.

Du ventricule gauche il est éjecté par l'intermédiaire de l'orifice aortique dans l'aorte [4].

L'ensemble de ces étapes forme la circulation sanguine illustrée sur (**la figure 1.2**).

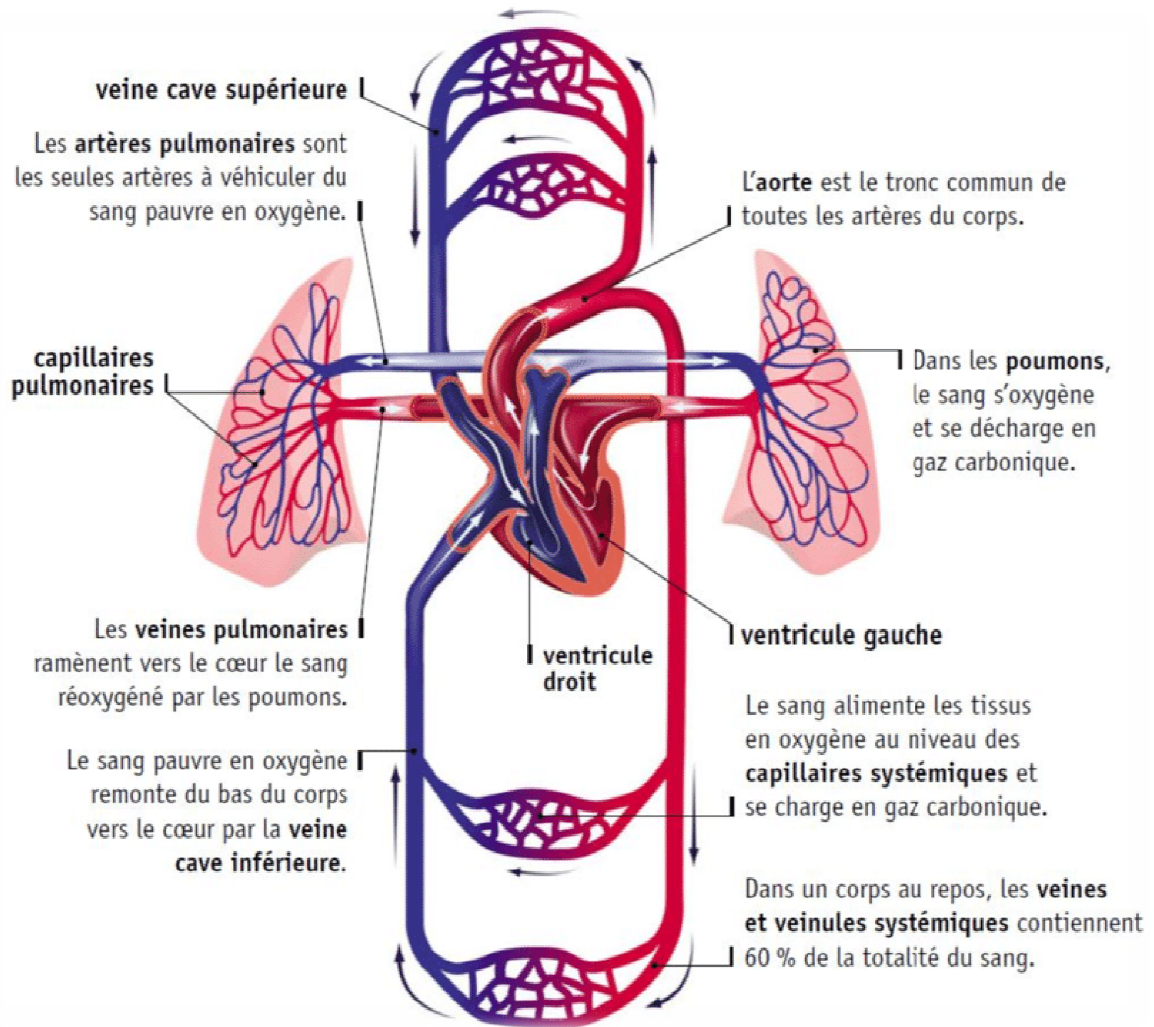


Figure 1.2. La circulation sanguine [4].

1.3.3 L'activité électrique du cœur

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. Dans le cœur, la dépolarisation prend normalement naissance dans le haut de l'oreillette droite (le sinus), et se propage ensuite dans les oreillettes,

Chapitre 1: Les notions de base de l'électrocardiographie

Induisant la systole auriculaire (**Figure 1.3**) qui est suivie d'une diastole (décontraction du muscle). L'impulsion électrique arrive alors au nœud auriculo-ventriculaire (AV), seul point de passage possible pour le courant électrique entre les oreillettes et les ventricules. Là, l'impulsion électrique subit une courte pause permettant au sang de pénétrer dans les ventricules. Elle emprunte alors le faisceau de His, qui est composé de deux branches principales allant chacune dans un ventricule. Les fibres constituant ce faisceau, complétées par les fibres de Purkinje, grâce à leur conduction rapide, propagent l'impulsion électrique en plusieurs points des ventricules, et permettent ainsi une dépolarisation quasi instantanée de l'ensemble du muscle ventriculaire, malgré sa taille importante, ce qui assure une efficacité optimale dans la propulsion du sang ; cette contraction constitue la phase de systole ventriculaire. Puis suit la diastole ventriculaire (décontraction du muscle) ; les fibres musculaires se re-polarisent et reviennent ainsi dans leur état initial [5].

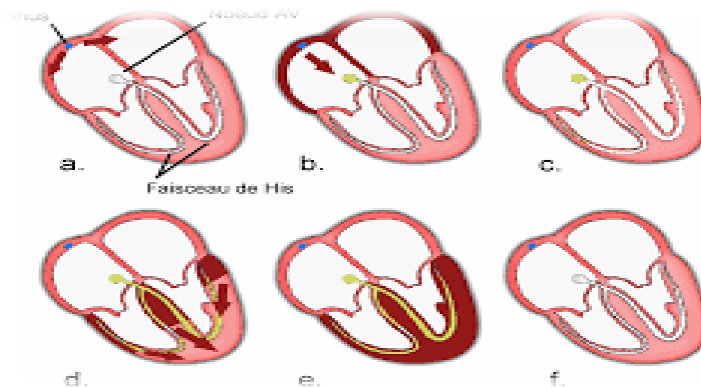


Figure 1.3. L'impulsion électrique se propage dans le muscle cardiaque et induit sa contraction. Elle prend naissance dans le sinus (a) puis se propage dans les oreillettes (b) entraînant leurs contractions (systole auriculaire). L'impulsion arrive alors au nœud auriculo-ventriculaire (AV) seul point 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 re-polarisent (f). Le cycle du battement cardiaque est alors terminé et le cœur est prêt pour un nouveau battement [5].

1.4 Le signal électrocardiogramme (ECG)

Le signal ECG (figure 1.4) regroupe un ensemble d'ondes connues par l'onde P, Q, R, S, et l'onde T qui sont en corrélation directe avec l'activité cardiaque. Ces ondes occupent des

Chapitre 1: Les notions de base de l'électrocardiographie

localisations temporelles traduisant des intervalles qui sont généralement différents pour un même signal sain ou présentant une pathologie cardiaque. Ces intervalles connus par l'intervalle RR, QT, PQ, ou encore le segment ST et le complexe QRS sont d'un intérêt diagnostique certain et varient de différentes manières dans différentes pathologies. Ils permettent d'évaluer l'état normal ou anormal de la propagation de l'évènement électrique [6]

Ainsi ces différents ondes et intervalles induire respectivement :

- 1) L'onde P : la déflexion correspondant à la dépolarisation des oreillettes droite et gauche. Sa durée normale est inférieure ou égale à 0.12s.
- 2) Le complexe QRS correspond à un ensemble de déflexions dues à la dépolarisation des ventricules, l'onde Q est la première : c'est une onde dirigé vers le bas, qui n'est pas toujours visible sur le tracé ; la seconde est l'onde R : elle est de grande amplitude et dirigée vers le haut ; la dernière est dirigée vers le bas : c'est l'onde S.
- 3) L'onde T : la déflexion correspondant à la repolarisation ventriculaire. C'est un phénomène purement électrique et pendant cette phase le cœur est mécaniquement inactif [8].

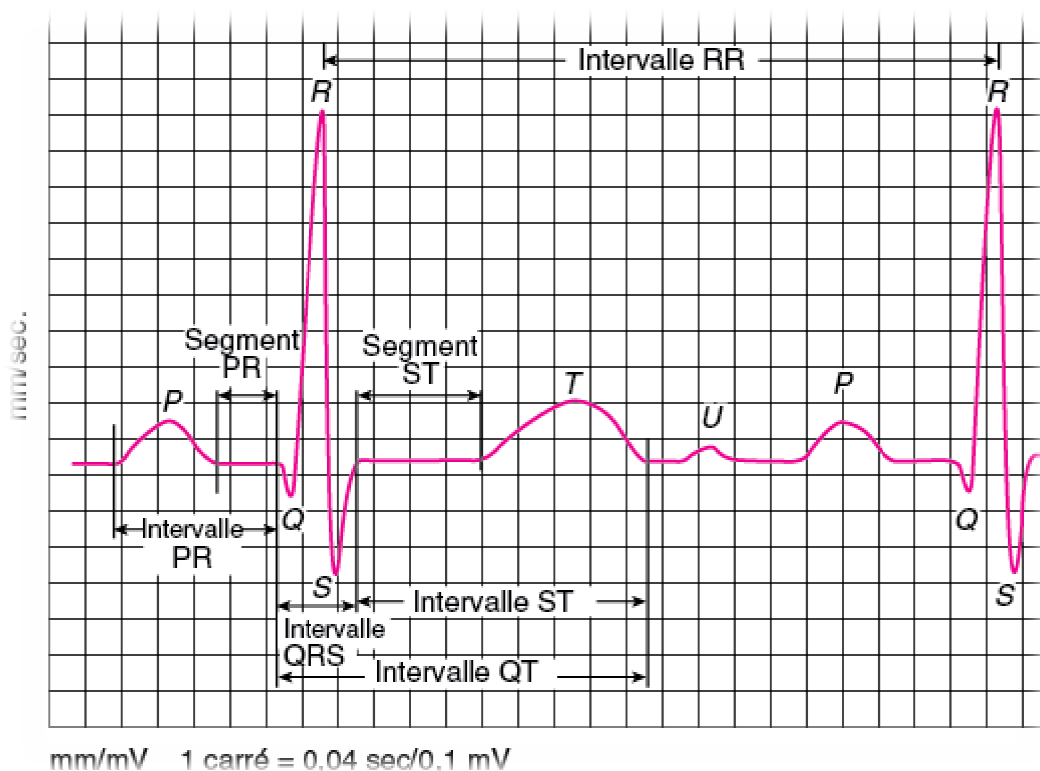


Figure 1.4. Le signal ECG [6].

Chapitre 1: Les notions de base de l'électrocardiographie

- L'intervalle P-R (ou P-Q) : correspond au délai entre la dépolarisation de l'oreillette et celle du ventricule. C'est le temps de propagation de l'onde de dépolarisation à travers les oreillettes, le nœud atrio-ventriculaire, le faisceau de His et le réseau de Purkinje, jusqu'aux cellules myocardiques ventriculaires.
- L'intervalle Q-T correspond au temps de systole ventriculaire qui va du début de l'excitation des ventricules jusqu'à la fin de leur relaxation.
- Le segment S-T correspond à la phase de repolarisation ventriculaire, phase durant laquelle les cellules ventriculaires sont toutes dépolarisées : il n'y a donc pas a priori de propagation électrique, le segment est alors isoélectrique.
- L'intervalle R-R sépare les sommets de deux ondes R successives et représente la fréquence instantanée. Ainsi, on parle de rythme sinusal régulier lorsque l'espace RR est régulier sur tout le tracé, avec des complexes QRS comparables.

L'activité du rythme cardiaque sur l'électrocardiogramme se fait donc en deux étapes vérifiant d'une part la régularité du rythme et d'autre part l'origine du rythme cardiaque qui peut être sinusal, jonctionnel, ventriculaire, ectopique, ou artificiel (pacemaker) [6].

1.4.1 Les dérivations électrocardiographiques

L'ECG est enregistré par des électrodes placées sur la paroi thoracique et les membres. Il existe trois types de dérivations: bipolaires, unipolaires périphériques et unipolaires précordiales.

1.4.1.1 Les dérivations périphériques :

Les dérivations périphériques (ou dérivations des membres) permettent d'étudier l'activité électrique du cœur sur le plan frontal. Elles sont obtenues au moyen de 4 électrodes appliquées au bras droit, au bras gauche, à la jambe gauche, et l'électrode de la jambe droite étant une électrode neutre destinée à éliminer les parasites électriques [7].

Les dérivations périphériques bipolaires (DI, DII, DIII) Ce sont les dérivations cardiaques classiques de l'électrocardiogramme décrites par Einthoven. Elles enregistrent la différence de potentiel entre deux électrodes placées à des extrémités différentes. Ces dérivations utilisent trois électrodes placées sur le sujet. Les électrodes sont placées sur les bras droit et gauche et sur la jambe gauche pour former un triangle (triangle d'Einthoven). Ces dérivations sont dites bipolaires parce qu'elles mesurent une différence de potentiel entre deux électrodes. Chaque

Chapitre 1: Les notions de base de l'électrocardiographie

côté du triangle formé par les trois électrodes représente une dérivation en utilisant une paire d'électrodes différente pour chacune des dérivations [7]. Les trois dérivations sont:

- DI (dérivation I) avec $DI = VL - VR$
- DII (dérivation II) avec $DII = VF - VR$
- DIII (dérivation III) avec $DIII = VF - VL$

Avec : VL : le potentiel sur le bras gauche, VR : le potentiel sur le bras droit et VF : le potentiel sur la jambe gauche

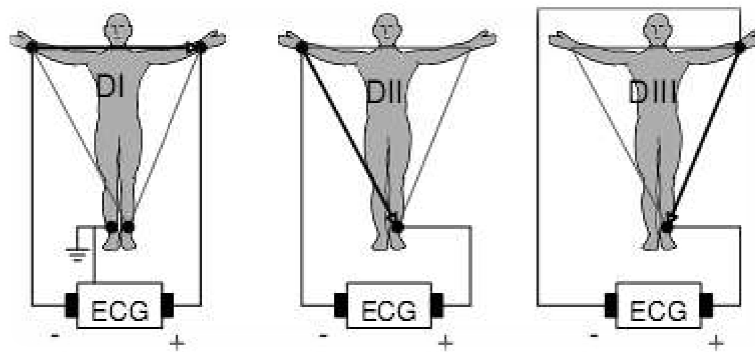


Figure 1.5. Montage d'Einthoven pour l'enregistrement des dérivations bipolaire des membres [8].

Les dérivations unipolaires des extrémités enregistrent la différence de potentiel entre un point théorique au centre du Triangle de Einthoven, ayant une valeur de 0 et l'électrode de chaque extrémité, permettant ainsi de connaître le potentiel absolu dans ladite électrode.

Dans un premier temps, on a nommé ces dérivations VR, VL et VF. Le V signifiant vecteur et R, L, F : droite, gauche et pied (en anglais). Plus tard, on a ajouté le 'a' minuscule qui signifie amplifiée (les dérivations unipolaires actuelles sont amplifiées par rapport aux premières) [7].

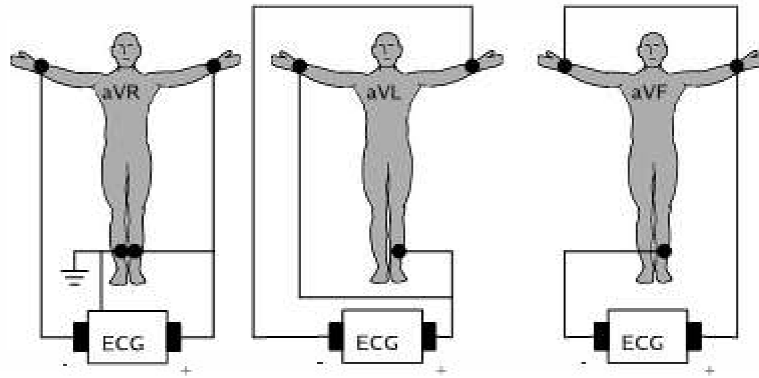


Figure 1.6. Montage d'Einthoven pour l'enregistrement des dérivations unipolaires des membres [8].

1.4.1.2 Les dérivations précordiales :

Il y a six pistes précordiales. Ils sont désignés par un V majuscule et un nombre compris entre 1 et 6. Ce sont des dérivations unipolaires, ils enregistrent le potentiel absolu du point où est placée l'électrode de même nom. Ce sont les dérivations les mieux adaptées pour dépister les anomalies du ventricule gauche, en particulier celles des parois antérieure et postérieure. Sur un électrocardiogramme normal, les QRS sont principalement négatifs dans les dérivations V1 et V2 et majoritairement positifs dans V4 à V6 (profil Rs) [7].

Ces dérivations sont positionnées comme suit :

- V1 : 4ème espace intercostal, bord droit du sternum (ligne parasternale).
- V2 : 4ème espace intercostal, bord gauche du sternum (ligne parasternale).
- V3 : à mi-distance entre V2 et V4.
- V4 : 5ème espace intercostal, ligne médio-claviculaire gauche.
- V5 : à mi-distance entre V4 et V6, sur la ligne axillaire antérieure.
- V6 : même niveau horizontal que V4 et V5, ligne axillaire moyenne.

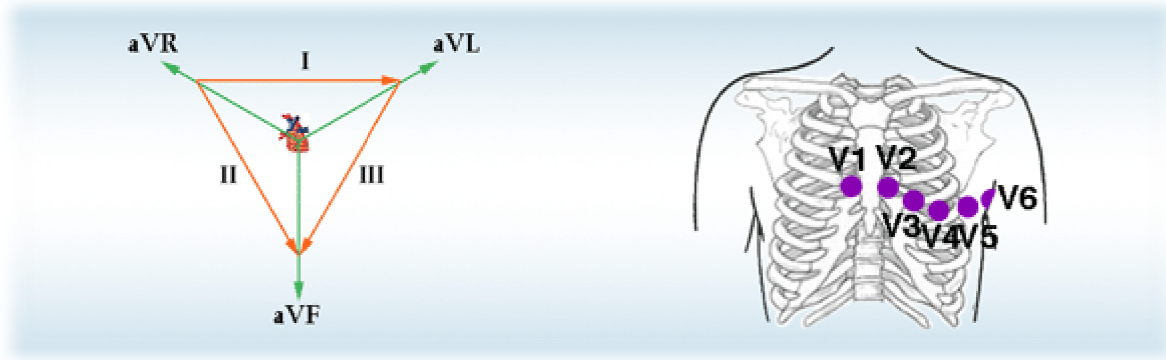


Figure 1.7. Dérivations thoraciques [9].

1.5 Les arythmies cardiaques

L'arythmie c'est une anomalie qui affecte la fréquence cardiaque normale.

1.5.1 La bradycardie

La bradycardie est caractérisée par une fréquence cardiaque inférieure à 60 bpm. Elle est dite d'origine sinusale, jonctionnelle ou ventriculaire, selon le site d'initiation de l'impulsion électrique à l'origine des battements considérés. Nous distinguons alors les cas suivants :

❖ Bradycardie sinusale

La bradycardie sinusale est caractérisée par la présence systématique d'une onde P avant les complexes QRS (Figure I.8). Le rythme est de l'ordre de 40 bpm.

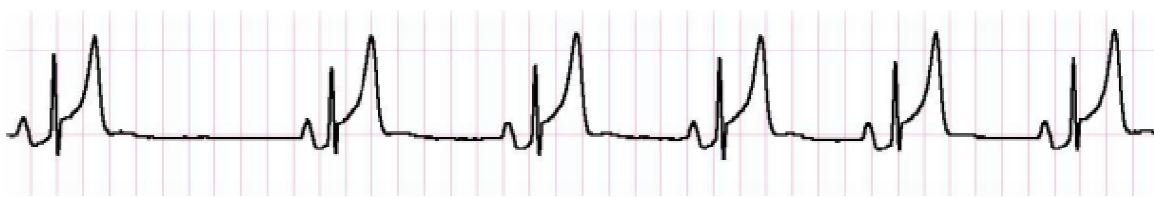


Figure 1.8. Bradycardie sinusale [9].

❖ Bradycardie d'origine jonctionnelle

L'impulsion électrique n'est plus transmise des oreillettes aux ventricules. La fréquence est de l'ordre de 30 à 60 bpm.

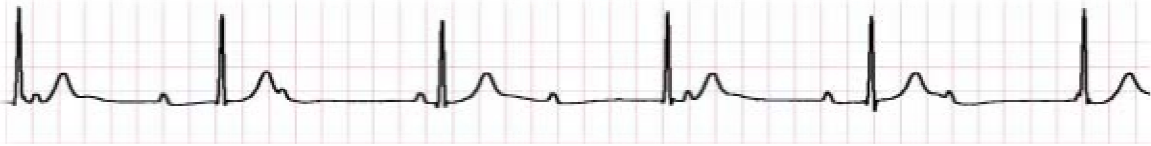


Figure 1.9. Rythme d'échappement jonctionnel à 35 bpm [10]

❖ Bradycardie d'origine ventriculaire

Le rythme a une fréquence très lente, entre 15 et 40 bpm.

1.5.2 La tachycardie

La tachycardie À l'inverse de la bradycardie, la tachycardie est caractérisée par une fréquence supérieure à 100 bpm, elle peut être d'origine sinusale, auriculaire ou ventriculaire. Dans le cas de la Tachycardie sinusale par exemple la fréquence est comprise entre 100 et 180 bpm et nous observons avant chaque complexe QRS, une onde P [10].

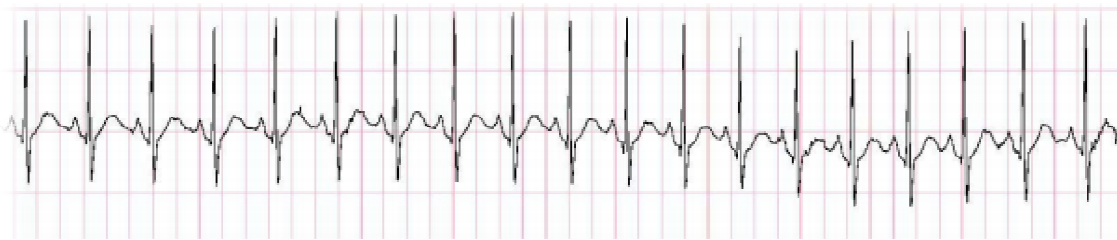


Figure 1.10. Tachycardie sinusale [10].

1.5.3 L'extrasystole

L'extrasystole est un battement cardiaque prématuré ou un surplus qui est ressenti comme un battement erratique ou manquant, comme si le cœur sautait un tour. C'est l'arythmie la plus fréquente. Nombreuses sont les personnes qui ont des extrasystoles et qu'elles ne perçoivent même pas. Si celles-ci ne s'accompagnent pas d'autres symptômes, ce qui constitue la majorité des cas, cette anomalie est bénigne et peut se produire dans un cœur sain. Il arrive parfois que l'extrasystole s'accompagne d'un bref étourdissement, mais c'est sans gravité [11].

1.5.4 La fibrillation auriculaire

Une fibrillation auriculaire est définie comme la contraction anarchique des oreillettes, qui entraîne une contraction rapide et irrégulière des ventricules. Il s'agit d'un trouble du rythme du cœur très fréquent, en particulier chez les sujets de plus de 60 ans. Le diagnostic est difficile car les symptômes de fibrillation auriculaire ne sont absolument pas ressentis [12].

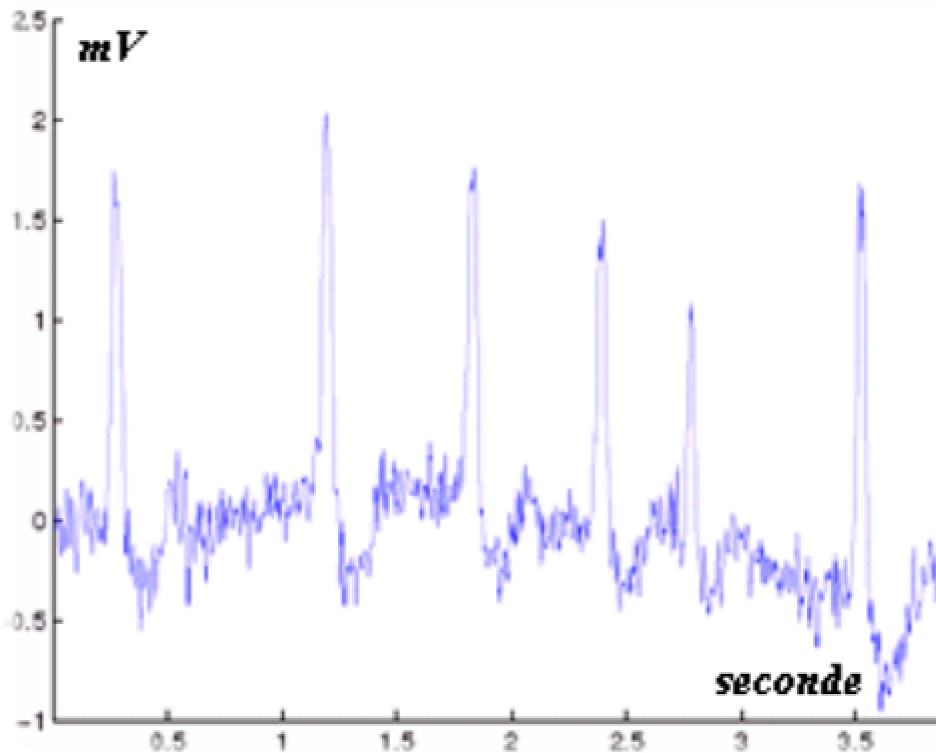


Figure 1.11. Fibrillation auriculaire [12].

1.5.5 Fibrillation ventriculaire

La fibrillation ventriculaire (FV) est parfois inaugurale (mort subite), précoce (fibrillation ventriculaire primaire de bon pronostic), parfois tardive (fibrillation ventriculaire secondaire) accompagnant alors un anévrisme du ventricule gauche suite à un mauvais diagnostic. La fibrillation ne peut pas se produire dans un milieu homogène. En fait, la période réfractaire, l'excitabilité, ainsi que la vitesse de conduction n'ont pas des propriétés constantes partout dans le tissu ventriculaire [12].

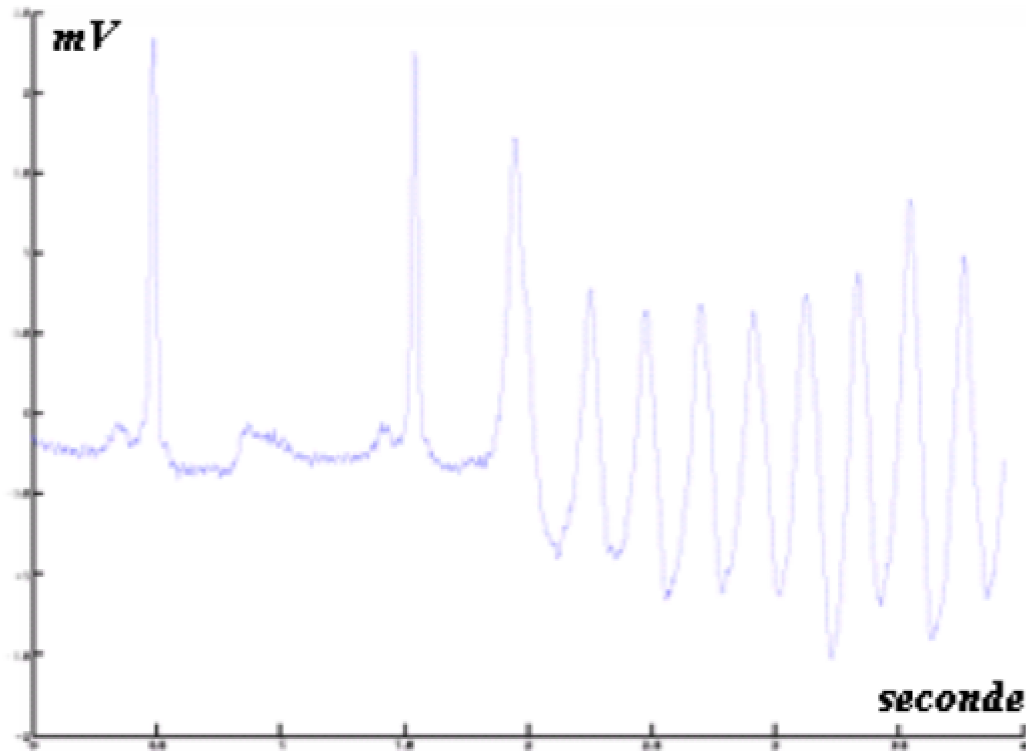


Figure 1.12. Fibrillation ventriculaire [12].

1.5.6 Bloc de branche

L'influx électrique une fois qu'il atteint le nœud auriculo-ventriculaire, il emprunte le faisceau de HIS ce dernier se scinde en deux branches, une branche droite qui véhicule l'influx dans le ventricule droit et une branche gauche véhicule l'influx dans le ventricule gauche qui se divise elle-même en deux hémibranches antérieure et postérieure. Le bloc de branche peut se manifester en tant que ralentissement de la conduction électrique dans les branches ou carrément le blocage de cette dernière. En effet si l'influx ne passe pas dans les branches alors la dépolarisation s'effectuera de proche en proche entre les cellules même du myocarde, ce qui se traduit par une dépolarisation des ventricules sur une durée plus longue, engendrant ainsi un QRS plus large sur le tracé de l'ECG [13]. On distingue les types suivants :

1.5.6.1 Bloc de branche droit

Le trouble de la conduction touche la branche droite, dans ce cas le ventricule droit est dépolarisé après le ventricule gauche, ceci se traduit alors sur le tracé de l'ECG par un

Chapitre 1: Les notions de base de l'électrocardiographie

complexe rSR' large en V1, un complexe qRS en dérivation V6 et une onde T inversée sur les dérivations V1, V2 et V3 (voire la figure 1.13).

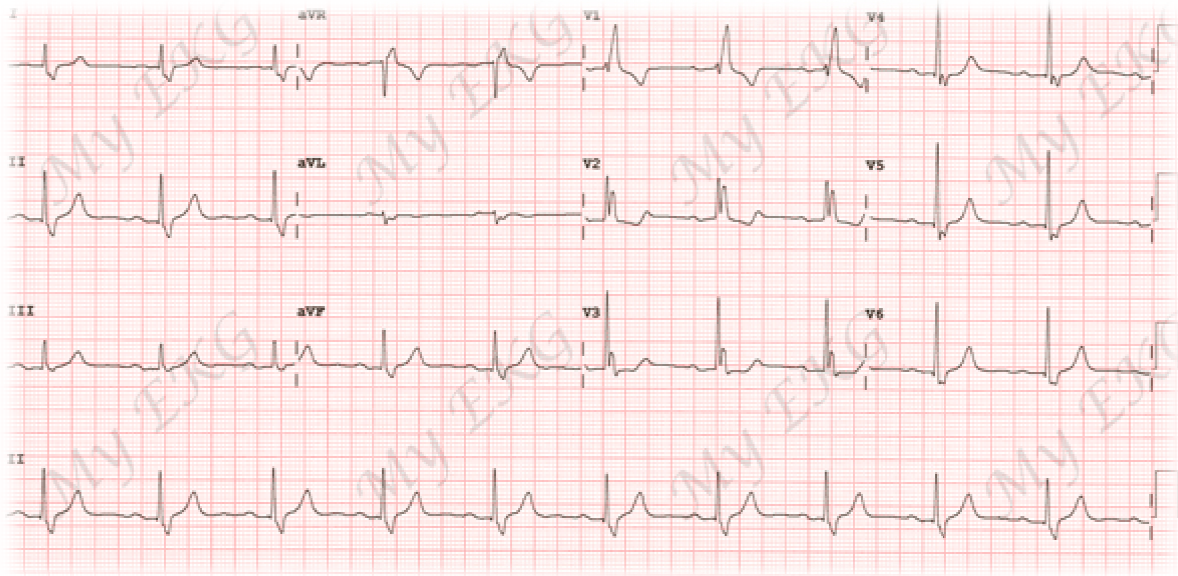


Figure 1.13. Bloc de branche droit représenté par les 12dérivtions [13].

1.5.6.2 Bloc de branche gauche

C'est la branche gauche qu'est concernée par le problème de conduction, le ventricule gauche est dépolarisé après le ventricule droit, ceci se traduit sur le tracé de l'ECG par un complexe QS large en V1, une onde R large en V6 avec une légère entaille sur son pic, et une onde T inversée en V5et V6 (voire la figure 1.14) [13].

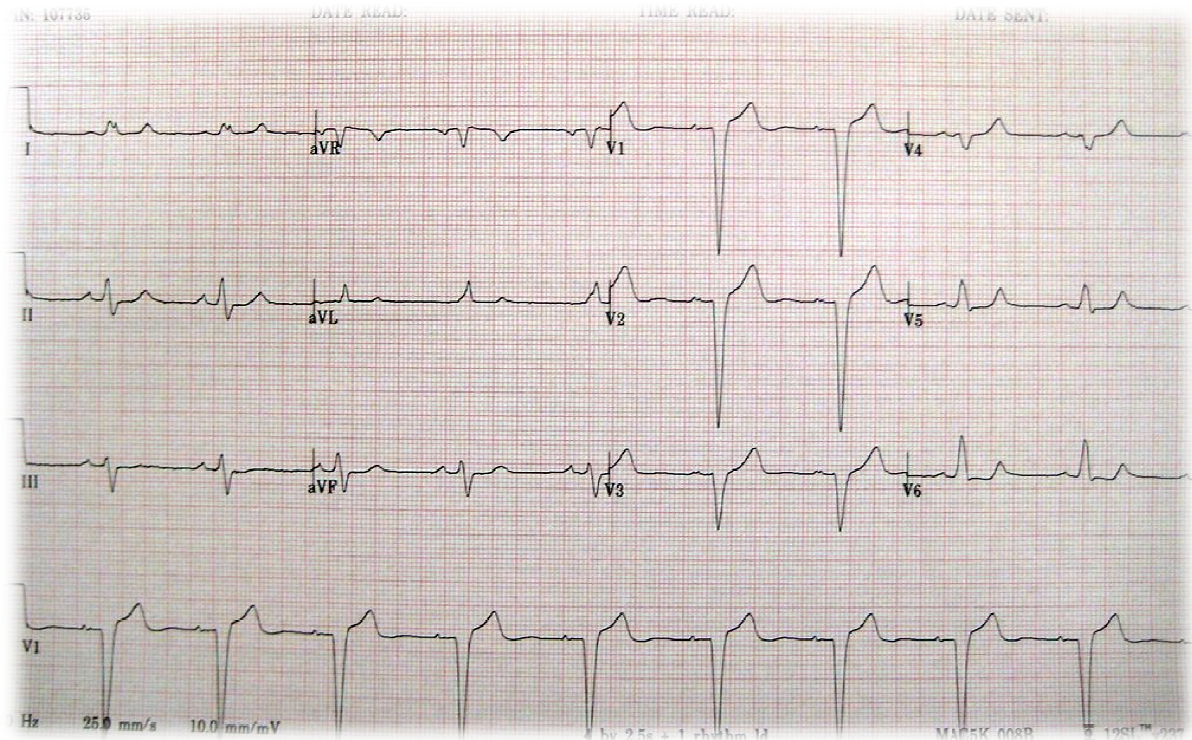


Figure 1.14. Bloc de branche gauche représenté par les 12 dérivation [13].

1.6 Artefacts visibles sur l'électrocardiogramme

Sur tout enregistrement électrocardiographique il peut apparaître des événements indésirables pouvant brouiller le tracé et, parfois, induire en erreur le diagnostic final. Ces bruits sont reconnaissables par l'œil expérimenté qui les identifie avant d'effectuer son diagnostic. Les effets indésirables peuvent avoir plusieurs sources: techniques, physiques ou pathologiques. Nous allons surtout développer l'aspect technique et physique des bruits et artefacts présents sur les tracés électrocardiographiques. En partant du principe que les bruits fréquents en électrocardiographie sont des bruits additifs, les caractéristiques de ces bruits auxquelles nous allons nous attacher sont l'amplitude, la périodicité et la bande spectrale. Les artefacts prennent une place particulière puisqu'ils induisent des modifications des performances des algorithmes d'analyse automatique de l'électrocardiogramme [14].

1.6.1. Bruits techniques

Le matériel utilisé lors de l'enregistrement doit être manipulé avec précaution car il peut être source de bruits lors de l'enregistrement. Les plus courants sont présentes ci-dessous.

✚ Bruit dû au secteur

Le réseau de distribution électrique peut parfois brouiller le signal électrocardiographique avec une onde dont l'harmonique principale est à 50 Hz comme montré sur la figure (1.15). Ce type de bruit apparaît sur tout l'enregistrement et peut être assez fort mais il s'élimine facilement avec un filtre sélectif car c'est un bruit haute fréquence a bande étroite.

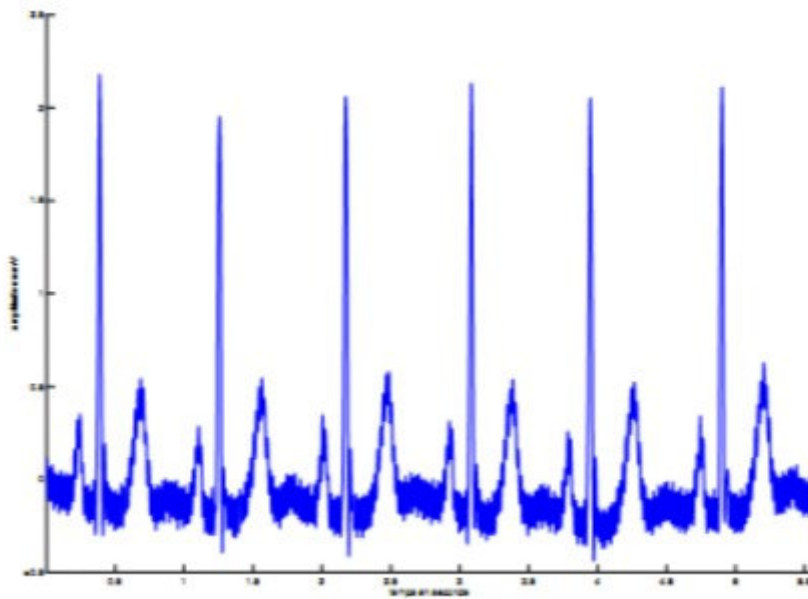


Figure 1.15. Interférence secteur 50 Hz [15].

✚ Bruit dû aux mouvements d'électrodes

Lorsque les électrodes sont connectées incorrectement, des sauts brusques de la ligne de base apparaissent. L'effet sur le tracé peut aller de la simple diminution d'amplitude à l'apparition de pics lorsque les électrodes sont en contact intermittent avec la peau. Ces pics peuvent parfois être confondus avec les ondes du tracé normal comme montré sur la figure (1.16). Ce type de bruit intermittent à bande spectrale large s'élimine difficilement car son énergie se trouve dans la même gamme de fréquence que le complexe QRS [14].

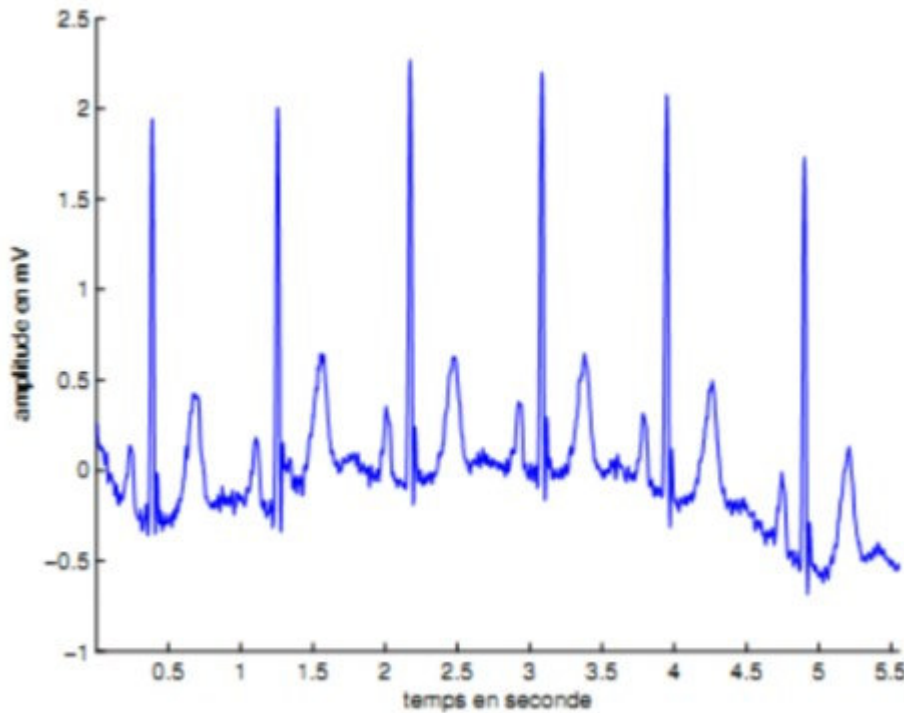


Figure 1.16. Bruit dû aux mouvements des électrodes [15].

✚ Autres bruits courants

Parmi les bruits courants on peut citer les artefacts dus aux mouvements des câbles électriques, la saturation des instruments de mesure, les mauvais câblages, les artefacts dus au port de vêtements synthétiques, etc.

1.6.2 Artefacts physiques

Les artefacts physiques sont dus aux activités électriques du corps humain telles que les commandes de contraction des muscles ou la respiration.

✚ Mouvements de la ligne de base

Lors de l'enregistrement de l'électrocardiogramme, l'activité respiratoire peut faire osciller la ligne de base de l'ECG à un rythme régulier comme montré sur la figure (1.17). D'autres perturbations peuvent avoir pour effet de déplacer temporairement la ligne de base comme, par exemple, les mauvais contacts entre la peau et les électrodes. Ces perturbations sont généralement peu gênantes pour l'analyse de l'ECG et peuvent être en grande partie filtrées car leur énergie se situe dans une bande de fréquence basse, qui empiète peu sur celle de l'ECG normal.

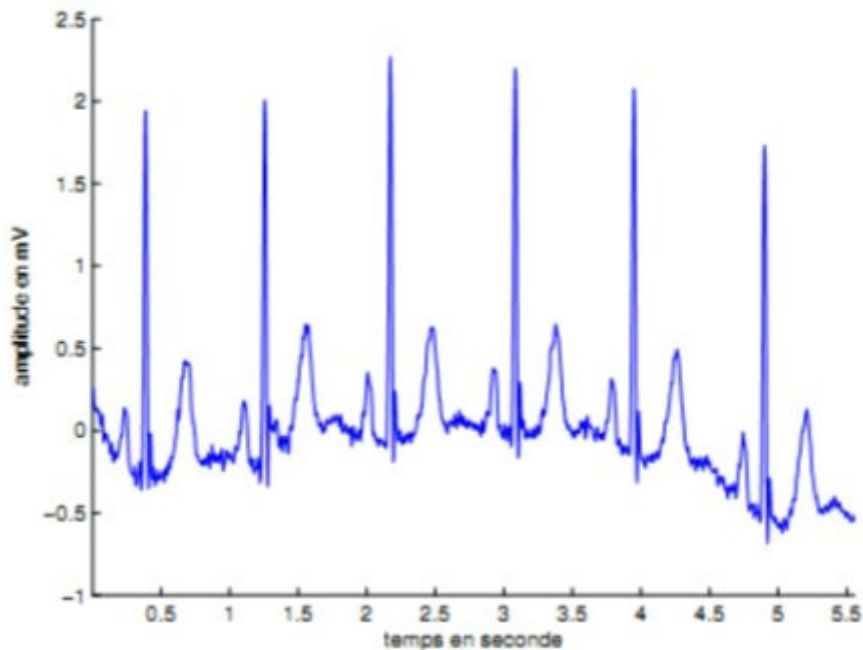


Figure 1.17. Dérive de la ligne de base [15].

Bruit myoélectrique ou tremblement somatique

La contraction d'un muscle est commandée par une dépolarisation des cellules musculaires et bien que les électrocardiographes soient construits pour être surtout sensibles aux fréquences du myocarde, l'ECG enregistre les contractions des muscles squelettiques comme montré sur la figure (1.18). L'aspect le plus courant est une oscillation à haute fréquence liée à la tension musculaire d'un sujet qui n'est pas convenablement détendu. Ces perturbations sont assez gênantes lorsque le patient bouge beaucoup ou lorsqu'il frissonne, elles peuvent noyer les ondes P et T et empêcher un diagnostic fiable. L'apparition de ces perturbations dépend de l'état du patient, s'il est très tendu ou atteint de maladie de Parkinson, l'enregistrement peut être de mauvaise qualité sur toutes les voies de l'ECG [14].

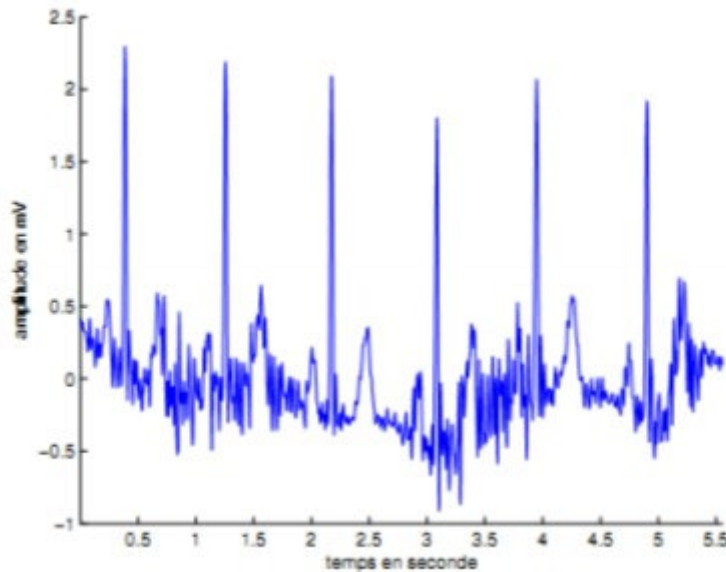


Figure 1.18. Bruit myoélectrique ou tremblement somatique [15].

1.7. La base de donnée MIT/BIH

La source de l'ECG inclus dans la base de données d'arythmies MIT/BIH est un ensemble de 48 enregistrements échantillonnés à une fréquence de 360 Hz. Chaque enregistrement comporte deux dérivations différentes du signal ECG. La base de données contient 23 enregistrements numérotés entre 100 et 124 pour le premier groupe; et de 25 enregistrements numérotés entre 200 et 234 pour le deuxième groupe. La durée de chacun des 48 enregistrements est de 30 minutes. Le premier groupe est prévu pour servir d'échantillon représentatif de variété de formes d'ondes qu'un détecteur d'arythmie pourrait rencontrer dans l'utilisation clinique courante; tandis que le deuxième groupe est choisi pour inclure une variété de cas pathologiques. Les sujets étaient 25 hommes âgés de 32 à 89 ans, et 22 femmes âgées de 23 à 89 ans. Les enregistrements de MIT/BIH sont tous annotés; ce qui veut dire que chaque battement (complexe QRS) est décrit par une étiquette (annotation). Typiquement, un fichier d'annotation pour un enregistrement MIT/BIH contient environ 2000 annotations [16].

Conclusion

Ce chapitre est une introduction aux éléments de base de l'électrophysiologie du cœur. Après une introduction sur l'anatomie du cœur humain, nous avons décrit brièvement l'activité électrique du cœur. Ces notions sont indispensables pour la bonne compréhension des problèmes étudiés par la suite. Dans une deuxième partie de ce chapitre nous avons exposé les différentes notions de l'électrocardiographie. Quelques pathologies cardiaques typiques ont été présentées et illustrées chacune par un enregistrement ECG correspondant. Nous avons terminé ce chapitre par donner les différents types d'artefacts qui peuvent affecter le signal ECG enregistré et une description de la base de données MIT/BIH.

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

2.1 Introduction

Une des méthodes les plus courantes pour obtenir des informations sur le système cardiovasculaire est d'étudier les caractéristiques des signaux électrocardiographiques enregistrés à travers la modélisation de signal ECG et ses différents paramètres et intervalles. En effet cette dernière est largement utilisée en génie biomédical [1].

Dans ce chapitre, nous allons d'abord décrire les différentes méthodes de filtrage de signal ECG, puis les méthodes de détection des ondes P,QRS, et l'onde T, en suite nous allons décrire les différentes méthodes d'analyse paramétriques les plus couramment utilisés dans le domaine biomédical, qui sont les modèles AR , MA et ARMA. Dans ce chapitre, le but n'est pas limite seulement à une meilleure identification, mais aussi le maintien de l'analyse spectrale en utilisant les paramètres extraits de chaque signal.

2.2 Les méthodes de filtrage de signal ECG :

Pour pouvoir segmenter efficacement les battements cardiaques, et ce sans altérer l'information clinique, un certain nombre de prétraitements sont nécessaires. La finalité de cette étape est d'atténuer, ou au mieux d'éliminer, les bruits présents dans le signal ECG brut tels que la dérive de la ligne de base ou les interférences du secteur à 50 Hz. Malheureusement, le bruit causé par l'activité musculaire est beaucoup plus difficile à filtrer à cause du chevauchement entre le spectre du bruit et celui de l'ECG [17]. Nous introduirons dans cette partie quelques méthodes clés de la littérature.

2.2.1. Filtrage de la dérive de la ligne de base

Pour l'analyse d'un enregistrement ECG, un œil exercé fait abstraction de cette ligne : elle est prise comme référence pour étudier la forme et la hauteur des différentes ondes cardiaques; néanmoins, dans l'objectif d'un traitement automatique d'un tel signal, il est impératif de la repérer précisément pour fixer le "zéro". L'élimination de cette dérive de ligne de base est nécessaire pour limiter les déformations morphologiques des ondes de l'ECG. Plusieurs types de méthodes ont été présentés dans la littérature afin d'éliminer ce type de bruit. Les méthodes s'appuient souvent sur un filtrage passe-haut, généralement basé sur des filtres à réponse impulsionnelle finie [18,19]. L'inconvénient majeur de ces méthodes est la distorsion du signal due au chevauchement des spectres de l'ECG et de la ligne de base. Il s'avère impossible de supprimer complètement ce type de bruit avec ces méthodes de filtrage

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

sans provoquer une distorsion du signal [20]. La fréquence de coupure du filtre doit être par conséquent un compromis entre l'élimination de la dérive de ligne de base et la déformation du signal : en général on choisit alors une fréquence de coupure légèrement plus basse que la fréquence de battement cardiaque la plus faible [21]. Afin de résoudre ces problèmes de chevauchement de spectres, un filtre avec des caractéristiques variantes dans le temps est proposé dans [22,23]; le filtre est implémenté comme un banc de filtres linéaires passe-bas où les fréquences de coupure sont alors contrôlées par les propriétés basses fréquences du signal ECG. Cette méthode nécessite la connaissance des instants d'apparition des complexes QRS au préalable. Dans la dérive de la ligne de base est estimée à partir d'un modèle polynomial. Le filtrage de la ligne de base peut être basé sur des splines [17].

2.2.2. Filtrage du bruit d'interférence à 50 Hz (ou 60 Hz) :

Ce type d'interférence est caractérisé par un signal de type sinusoïdal de 50 Hz (voire 60 Hz dans d'autres pays) généralement accompagné de quelques harmoniques. Une méthode simple permettant de réduire ce type de bruit consiste à éliminer une fréquence particulière ou une gamme de composantes fréquentielles; le filtre qui convient alors est appelé filtre Notch ou filtre à bande étroite. Lorsque la fréquence des perturbations n'est pas stable sur 50 Hz, la variation de fréquence, considérée de type gaussienne, n'est pas toujours centrée sur 50 Hz. Pour remédier à ce type de problème, des méthodes basées sur du filtrage adaptatif ont également été proposées [24].

2.2.3. Filtrage du bruit de l'électromyogramme :

Le bruit causé par l'activité musculaire est beaucoup plus difficile à filtrer. La méthode de filtrage la plus utilisée est basée sur le moyennage sur plusieurs battements successifs du signal. En supposant le bruit aléatoire et stationnaire, la réduction du bruit est alors proportionnelle à la racine carrée du nombre de battements utilisés pour ce moyennage; toutefois un moyennage sur un nombre trop important de battements peut causer des erreurs importantes à cause de la dispersion du signal entre les battements (notamment à l'effort). Ce type de filtre pour atténuer le bruit électromyographique n'est pas applicable lorsqu'on cherche des informations sur les ondes battement par battement [17].

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

2.3. Détection des ondes P, R, Q, S, et T

2.3.1. Filtrage de signal ECG :

Généralement les ondes R ont une énergie maximale dans la bande passante 5-15 Hz [25]. Il est donc préférable de filtrer ce signal comme une première étape et avant toute analyse. Le filtre utilisé est un filtre numérique récursif de Butterworth de bande passante 5-15 Hz, et d'ordre 5 [26].

2.3.2. Détection des pics R :

Une fois le signal ECG est filtré, on procède à la détection des complexes QRS. L'algorithme utilisé [26] contient les quatre étapes décrites ci-dessous:

- Application de la prédiction linéaire et le calcul de l'erreur résiduelle

La prédiction linéaire permet de donner des informations au niveau de complexe QRS, compte tenu que l'erreur résiduelle contient beaucoup plus d'informations au niveau de la bande étroite du complexe QRS. En effet, le complexe QRS dure très peu de temps et ayant une amplitude très élevée par rapport au reste du signal, ce qui donne une erreur résiduelle plus élevée au niveau du complexe.

- Transformation non linéaire

La transformation non linéaire permet une bonne localisation des pics R. Pour cela il faut éliminer le signe et assurer la symétrie de détection entre les ondes R et les ondes Q et S. elle est donnée par l'équation (2.1):

$$y(n) = \begin{cases} |x(n)x(n-1)x(n-2)| & \text{si } x(n), x(n-1), x(n-2) \text{ ont le même signe} \\ 0 & \text{ailleurs} \end{cases} \quad (2.1)$$

Détection des pics R

Cette étape consiste en une recherche de maxima capable de ne pas prendre en considération les maxima trop faibles qui peuvent correspondre au bruit en lui affectant un seuillage adaptatif.

- Tout d'abord on fixe un seuil initial avec une valeur égale à 40% de l'amplitude maximale observée pendant les deux premières secondes du signal.

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

- Ensuite la valeur du seuil est divisée par deux.
- Une période réfractaire égale à 100 ms est imposée.
- S'il y aura un dépassement de seuil adaptatif, le seuil prendra la valeur de l'amplitude de signal. Sinon, si la période réfractaire est écoulée la valeur de seuil est divisée par deux.

Afin de détecter les ondes P il faut d'abord détecter les ondes Q, S, et T. Dans ce travail l'algorithme de Pablo-Laguna [27] a été utilisé. Il est décrit dans ce qui suit.

2.3.3. Détection des ondes Q et S :

Après la détection des pics R, on procède à la détection des ondes Q et S ou bien la détection de début et de la fin du complexe QRS noté bQRS et eQRS en utilisant l'algorithme de Laguna et al. Le signal ECG est dérivé par un filtre dérivateur I afin d'obtenir le maximum d'information au niveau du complexe QRS. La fonction de transfert en z de ce filtre est donnée par l'équation (2.2).

$$H_1(z) = \frac{Y(z)}{X(z)} = -2z^{-4} - z^{-3} + z^3 + 2z^4 \quad (2.2)$$

Après la première dérivé, le signal obtenu est appliqué à un deuxième dérivateur (dérivateur II) qui fait augmenter le gain du signal résultant. La fonction de transfert en z de ce filtre est donnée par l'équation (2.3) :

$$H_2(z) = \frac{-z^{-3} + z^3}{-z^{-1} + z} \quad (2.3)$$

Afin d'obtenir un maximum unique pour chaque début et fin de complexe QRS, un filtre récursif de Butterworth d'ordre 1 et de fréquence de coupure de 1 Hz est appliqué. Ce filtre présente un inconvénient de retard de phase mais il n'a aucune influence sur notre travail, en effet on a utilisé le même type de filtre. Le deuxième filtre dérivateur permet de préserver l'information correspondant au début et la fin du complexe QRS.

2.3.4. Détection des ondes T :

La détection de l'onde T est plus difficile que les autres ondes, non seulement à cause de sa morphologie variable (normale, inversé, trop large ou trop étroite avec une amplitude grande ou faible), mais aussi à cause de la variation instable du rythme cardiaque. Pour cela et à partir de la position de l'onde R, une fenêtre de largeur (B_{win} , E_{win}) est définie. Cette fenêtre est donnée par la relation suivante [27].

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

$$(B_{win}, E_{win}) = \begin{cases} (140,500RR_{av} > 700ms \\ (100,0.7 * RR)ms & RR_{av} < 700 ms \end{cases} \quad (2.4)$$

Pour éviter une fausse détection de l'onde T, la taille de la fenêtre est diminuée quand l'intervalle RR diminue. L'intervalle moyen de RR noté (RR_{av}) est utilisé pour éviter les grandes variations dans l'intervalle RR. Afin de détecter la fin de l'onde T, une deuxième dérivation est faite sur le signal représentant les ondes T, afin d'obtenir un maximum au niveau des déflexions qui représentent les bords des ondes T.

2.3.5. Détection des ondes P :

Une fois les ondes Q, R, S, T sont détectés, on procède à la détection des ondes P. pour cela on fait une soustraction entre le signal original et les signaux représentant les ondes Q, R, S, et T respectivement. Le signal résultant est dérivé par un filtre dérivateur $H_2(z)$ donné par l'équation (2.3) afin d'augmenter le gain de notre signal. Ensuite le signal obtenu est filtré par un filtre numérique de Butterworth d'ordre 4 pour diminuer le bruit restant et faire un lissage de notre signal [1].

2.4. Les différents types des modèles autorégressifs et linéaires :

Le principe d'une identification paramétrique ou non paramétrique est d'extraire un modèle mathématique à partir d'observations. Le modèle doit permettre de calculer la sortie du procédé et à n'importe quel instant t si les conditions initiales du système sont connues.

Pour cela on peut se servir des valeurs des entrées aux instants présents et précédents ($u(t)$, $u(t-1)$, ...) et des valeurs précédentes de la sortie ($y(t-1)$, $y(t-2)$, ...) dans le cas d'un modèle régressif. Il est tout de même important d'avoir des connaissances basiques du système pour choisir un type de modèle adapté. On trouve ainsi :

Il est tout de même important d'avoir des connaissances basiques du système pour choisir un type de modèle adapté. On trouve ainsi :

- Modèles possédant une entrée/une sortie (Single Input/ Single Output SISO) ou plusieurs entrées et plusieurs sorties (Multiple Input/Multiple Output MIMO)
- Modèles linéaires ou modèles non-linéaires (dans ce cas, qu'est-ce qui est non-linéaire et en fonction de quoi).

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

- Modèles continus ou modèles discrets.
- Modèles régressifs ou modèles indépendants : pour un modèle régressif, la sortie à un instant t , $y(t)$, dépend des instants précédents ($y(t-i)$).
- Modèles stochastiques ou modèles déterministes.

L'identification nécessite alors une structure de modèle connu a priori pour venir identifier dans cette structure différents paramètres. Le paragraphe suivant précise et décrit les différentes étapes dans la procédure d'identification [1].

2.5. Algorithmes d'identification paramétrique

L'algorithme général d'identification paramétrique comprend six étapes principales, comme il est illustré dans **la figure 2.1** [28]. Il s'agit de:

- La base de données
- Le Choix de la classe du modèle.
- l'estimation du modèle.
- la sélection d'ordre.
- la validation du modèle.
- et l'extraction des paramètres du modèle.

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

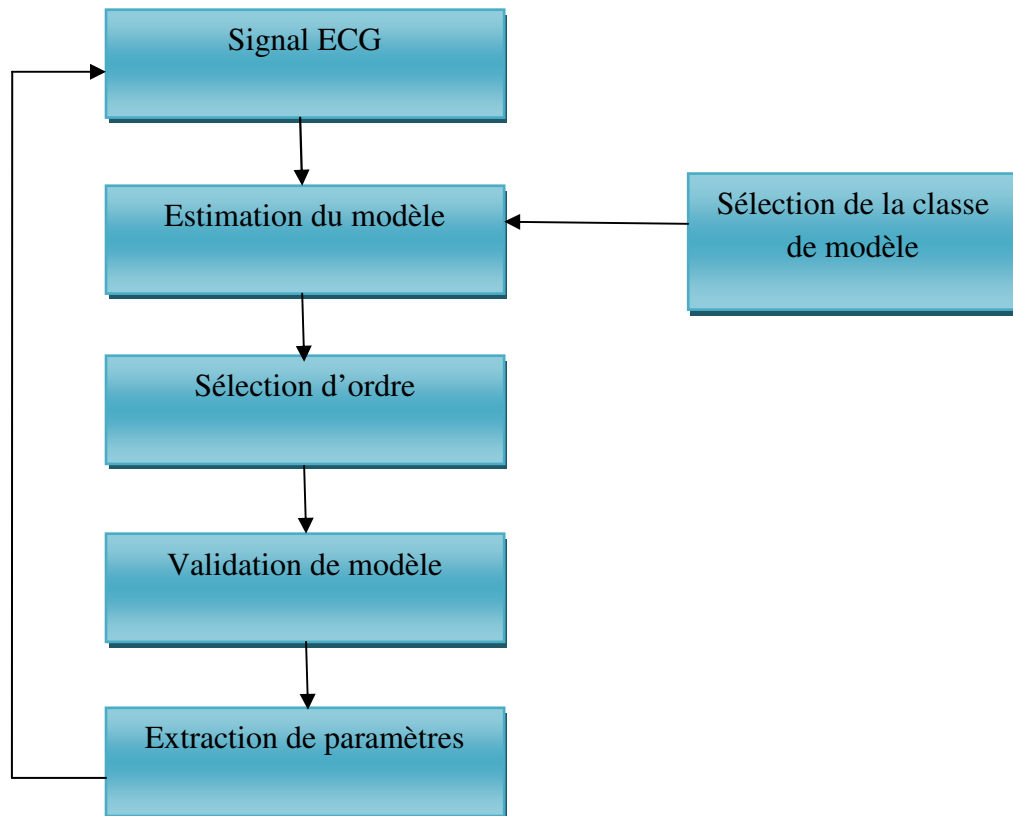


Figure 2.1. La procédure générale d'identification [1]

2.5.1. Sélection de la classe du modèle

Il y a une certaine confusion dans la sélection du type de modèle. Cette étape consiste à choisir le modèle le plus approprié. Le choix de la structure du modèle est basé sur la compréhension des systèmes physiques et dépend de la dynamique et des caractéristiques de bruit du système. En utilisant un modèle avec plus des paramètres n'est pas toujours mieux, car il peut conduire à une dynamique inexistante.

Trois types de modèles sont connus dans l'identification du système: le modèle de boîte noire, modèle de boîte grise (gray-box), et le modèle défini par l'utilisateur.

Le modèle de boîte noire suppose que les systèmes ne sont pas connus et tous les paramètres du modèle sont réglables sans tenir compte du contexte physique. En effet les paramètres ne peuvent pas être tous réglés arbitrairement.

Le modèle gray-box suppose qu'une partie de l'information sur les dynamiques sous-jacentes ou certains des paramètres physiques sont connus et les paramètres du modèle peuvent avoir certaines contraintes.

Le modèle défini par l'utilisateur suppose que les modèles paramétriques couramment utilisés ne peuvent pas représenter le modèle que nous voulons estimer.

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

2.5.2. Sélection de l'ordre

Une fois le modèle est choisi, on détermine son ordre. Cela signifie le nombre de coefficients qui permettent une meilleure modélisation de ce signal dans le cas d'un modèle paramétrique. L'ordre du modèle ne peut pas être déterminé avec précision, sans une analyse plus détaillée. Le choix d'ordre est en général obtenu en minimisant l'erreur entre la sortie et le signal d'entrée du modèle. Dans la pratique, deux critères sont utilisés pour déterminer l'ordre. Ces critères sont fondés sur des considérations statistiques, l'erreur de prédiction finale et le critère d'Akaike [29] :

a) Critères de l'erreur de prédiction finale (EPF):

Le critère de l'erreur de la prédiction final (EPF) fournit une mesure de la qualité du modèle en simulant la situation où le modèle est testé sur un ensemble de données différentes. Après que le calcul de plusieurs modèles différents soit fait, on peut les comparer en utilisant ce critère. Le modèle le plus précis sera donc le modèle générant la plus petite EPF.

EPF est donnée par l'équation (2.5):

$$EPF(k) = \sigma_k^2 \left(\frac{N+K+1}{N-K-1} \right) \quad (2.5)$$

Où σ_k^2 est la variance du bruit blanc pour l'ordre k.

N: est le nombre des échantillons du signal.

b) Critère d'information d'Akaike (AIC)

C'est une mesure de la qualité d'un modèle statistique. Il a été proposé par Hirotugu Akaike en 1973. En effet lorsqu'on estime un modèle statistique, il est possible d'augmenter la vraisemblance du modèle en ajoutant un paramètre. On choisit alors le modèle avec le critère d'information d'Akaike le plus faible. Ce critère est caractérisé par l'équation (2.6):

$$AIC(K) = N \ln(\sigma_K^2) + 2(K + 1) \quad (2.6)$$

La figure 2.2 montre un exemple d'optimisation de l'ordre d'identification pour un signal RR de la base de données MIT BIH. D'après les résultats trouvés, nous avons un ordre optimal de 2. D'autre part nous montrons que l'erreur décroît en fonction de l'ordre de prédiction ce qui est conforme à la littérature.

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

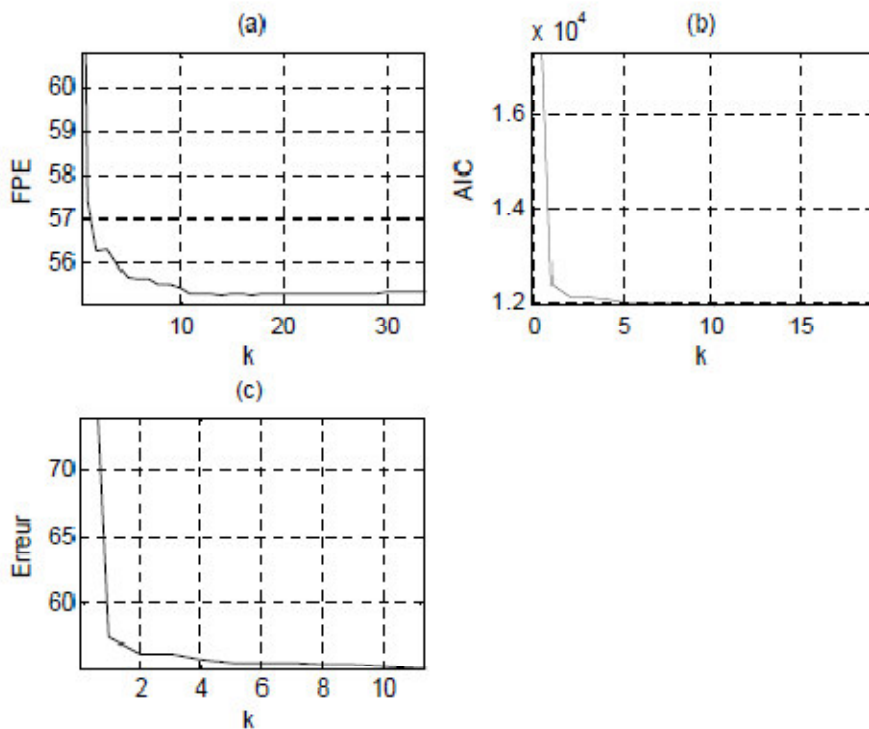


Figure 2.2. Optimisation de l'ordre de prédiction (a) FPE, (b), AIC, (c) évolution de l'erreur [1].

2.5.3. Procédure de validation du modèle

Elle se réfère à la détermination du modèle le plus efficace. Cette validation s'effectue en calculant le pourcentage d'identification du signal simulé et le signal original en utilisant L'équation (2.7) [30].

$$\text{fit} = 100 \left(1 - \frac{\text{norm}(y - \text{RR})}{\text{norm}(\text{RR} - \frac{1}{n} \sum_{i=0}^n \text{RR}_i)} \right) \quad (2.7)$$

Où y et RR' est la sortie (signal simulé) et RR est l'entrée (signal réel).

2.5.4. Extraction des paramètres du modèle

Cette étape consiste à la détermination des paramètres du modèle estimé; en utilisant le critère des moindres carrés (LS) pour le modèle AR et les moindres carrés généralisés (GLS). Pour les autres modèles à l'exception des modèles combinés où les paramètres ont été obtenus par la résolution de deux tâches d'identification distinctes [31].

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

2.6. Modèles paramétriques autorégressifs et linéaires

La structure générale du modèle linéaire simple est illustrée dans **la figure 2.3**. Elle offre une grande souplesse pour les deux parties, la dynamique du système et la dynamique stochastique.

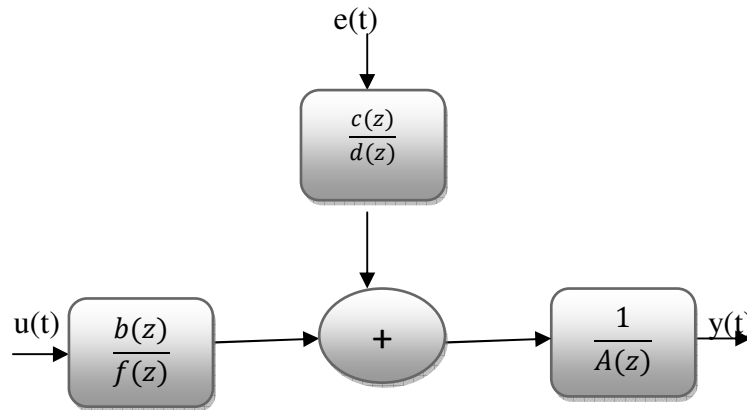


Figure. 2.3 Structure des modèles paramétriques autorégressifs et linéaires [16].

Les modèles paramétriques décrivent des systèmes en termes des équations différentielles et des fonctions de transfert en utilisant la transformée en z .

En général, on peut décrire le système en utilisant l'équation (2.8), qui est connu comme étant le modèle général polynomial linéaire ou le modèle général linéaire [28].

$$y(n) = z^{-k}G(z^{-1},\theta)u(n) + H(z^{-1},\theta)e(n) \quad (2.8)$$

Où $u(n)$ et $y(n)$: sont respectivement l'entrée et la sortie du système.

$e(n)$: est le bruit blanc.

$G(Z^{-1},\theta)$: est la fonction de transfert du système.

$H(Z^{-1},\theta)$: est la fonction de transfert de perturbation décrite comme un bruit blanc $e(t)$ filtré.

Dans le cas de cette étude $u(n)$ représente l'intervalle RR du signal ECG. L'équation (2.8) peut-être alors réécrite par l'expression suivante (équation 2.9)

$$A(z)RR'(n) = \frac{B(z)}{F(z)}RR(n) + \frac{C(z)}{D(z)}e(n) \quad (2.9)$$

$RR(n)$: est le nième intervalle RR.

n : est le nombre de battements.

$RR'(n)$: est le signal simulé (la sortie).

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

Des modèles plus simples peuvent être générés par la mise en œuvre d'un ou plusieurs polynômes $A(z)$, $B(z)$, $C(z)$, $D(z)$, ou $F(z)$ égal à 1. Elle conduit à des modèles tels qu'AR, MA, ARX, ARMA, OE, et BJ. Le choix de ces modèles est limité à un seul des cas particuliers illustres dans le tableau 2.1:

Coefficients	Modèle
A	AR
B	MA
A, B	ARX
A, B, C	ARMAX
B, F	OE (out put error)
B, F, C, D	BJ (Box Jenkins)

Tableau. 2.1 .Différents modèles autorégressifs et leurs coefficients [1].

Les polynômes $A(z)$, $B(z)$, $C(z)$, $D(z)$, $F(z)$ sont des filtres RIF d'ordre k [32].

$$A(z) = 1 + a_1z^{-1} + \dots + a_{k_a}z^{-k_a} \quad (2.10)$$

$$B(z) = b_0 + b_1z^{-1} + \dots + b_{k_b}z^{-k_b} \quad (2.11)$$

$$C(z) = c_0 + c_1z^{-1} + \dots + c_{k_c}z^{-k_c} \quad (2.12)$$

$$D(z) = 1 + d_1z^{-1} + \dots + d_{k_d}z^{-k_d} \quad (2.13)$$

$$F(z) = 1 + f_1z^{-1} + \dots + f_{k_f}z^{-k_f} \quad (2.14)$$

Une fois ces paramètres sont extraits, la densité spectrale de puissance (DSP) peut être estimée en utilisant ces paramètres. D'une manière générale elle est donnée par l'équation (2.15).

$$S(f) = T \partial_{RR}^2 \left| \frac{B(z)C(z)}{A(z)F(z)D(z)} \right|_{z=e^{j2\pi ft}}^2 \quad (2.15)$$

Dans le cas de ce mémoire T est la moyenne du rythme cardiaque.

2.6.1. Modèle Autorégressif (AR)

Le modèle autorégressif AR illustré dans la figure 2.4 permet une estimation simple et efficace du signal sans connaître la source et la modification qui l'affecte. Ce qui permet de

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

supposer qu'il a été généré par un bruit blanc qui passe à travers un filtre. Cette hypothèse permet la modélisation de n'importe quel signal sans restriction. [33].

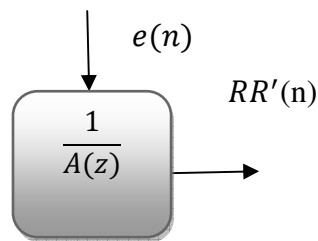


Figure 2.4. Structure du modèle AR [33].

2.6.2. Modèle à moyenne ajusté (Moving Average MA)

Un modèle moving average illustré sur la figure 2.5 est un filtre tout-zéro (composé uniquement de zéros) au travers duquel passe un bruit blanc. On détermine les coefficients de ce filtre tels que le signal à sa sortie soit le plus proche possible du signal à modéliser. L'équation d'identification des séries RR par le modèle MA d'ordre 2 est donnée par l'équation (2.16) [33].

$$RR'(n) = RR(n - 1) + 0.4519RR(n - 2) + 0.8778RR(n - 3) + e(n) \quad (2.16)$$

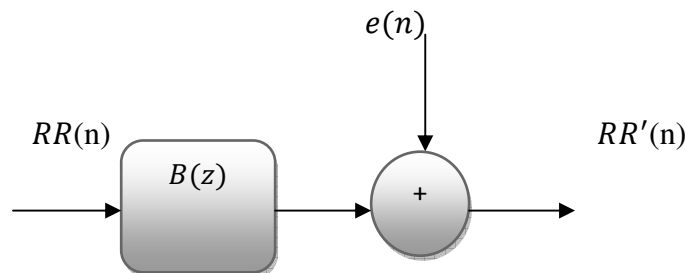


Figure 2.5. Structure du modèle MA [33].

2.6.4. Modèle ARMA

Le modèle ARMA (Autoregressive Moving Average) illustré dans la figure 2.6 utilise les attributs du modèle ARX mais inclut une fonction de transfert ($c(z)$) avec une moyenne du bruit blanc ajustable et présente aussi une sensibilité au bruit. Les modèles ARMAX sont utiles lorsque nous voulons dominer les perturbations qui entrent tôt dans le processus. Par exemple, un bruit respiratoire qui affecte le rythme cardiaque.

Chapitre 2 : Méthodes d'extraction de caractéristiques pour les signaux ECG (fonctions d'analyse et de modélisation)

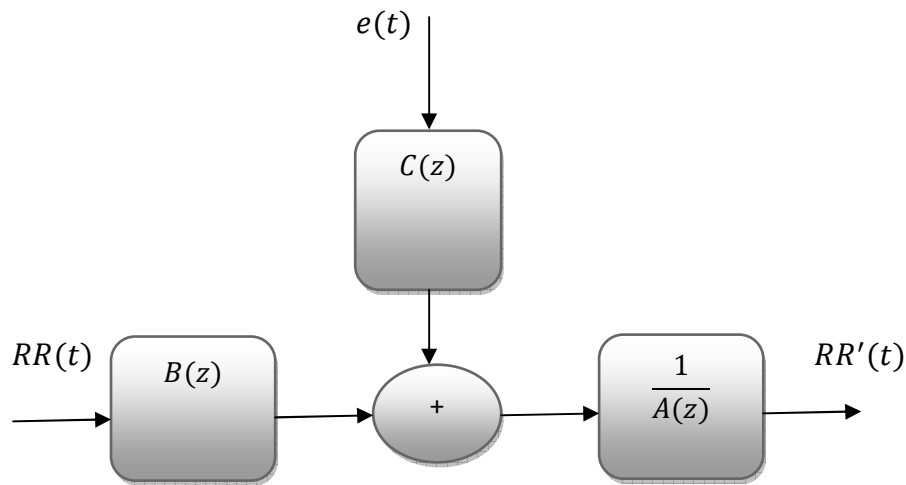


Figure 2.6. Structure du modèle ARMAX [33].

Conclusion

Dans ce chapitre nous avons présenté quelques méthodes de filtrage de signal ECG puis on a cité quelques méthodes permettent la détection des paramètres d'un signal ECG notamment le complexe QRS et les ondes T et P. Dans une deuxième partie de ce chapitre, une identification linéaire autorégressive et paramétrique a été présentée brièvement. Les modèles linéaires simples représenté dans ce chapitre sont : le modèle AR, MA et ARMA.

3.1 Introduction

En analysant l'ECG, il est possible de détecter la fréquence et la cohérence des battements cardiaques et d'identifier d'éventuelles irrégularités dans le fonctionnement du cœur. Les techniques d'extraction modernes jouent un rôle important dans diverses tâches médicales et techniques. Chaque méthode extrait un ensemble particulier de coefficients du signal ECG en fonction de sa nature. Ces coefficients, c'est-à-dire les caractéristiques extraites, sont utilisés pour effectuer et réaliser la classification du signal ECG pour la variété des tâches. Ce chapitre décrit un ensemble de techniques utilisées pour prétraiter les signaux ECG et extraire un ensemble de caractéristiques, de paramètres AR qui peuvent ensuite être utilisés pour caractériser le signal ECG. Les paramètres extraits de l'AR appliqués sont utilisés pour accomplir deux tâches, la classification ECG et les systèmes biométriques basés sur l'ECG. Les caractéristiques extraites appartenant à chaque signal ECG sont classées en groupes correspondants à trois affections cardiaques normales, à l'arythmie par fibrillation et à l'arythmie ventriculaire. Le système biométrique basé sur l'ECG utilise des paramètres AR extraits comme entrée du classificateur avec d'autres paramètres décrits dans ce chapitre, afin d'identifier les individus. La sélection et l'extraction des caractéristiques sont l'une des étapes cruciales pour des résultats de classification précis. Dans ce chapitre, les résultats obtenus en utilisant la méthode de modélisation AR sont comparés aux résultats obtenus en utilisant d'autres techniques d'extraction et les résultats pertinents discutés et les conclusions résumées.

3.2 Méthode d'extraction du système biométrique ECG

L'identification humaine basée sur la biométrie est essentiellement un problème de reconnaissance des patients qui implique des étapes de prétraitement, d'extraction de caractéristiques et de classification. **La figure 3.1** illustre le schéma fonctionnel de l'approche proposée.

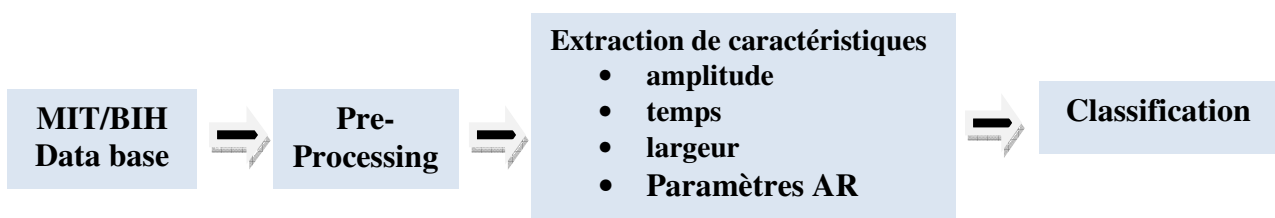


Figure 3.1: Schéma fonctionnel du système d'identification proposé basé sur l'ECG.

L'extraction des caractéristiques ECG joue un rôle important dans le diagnostic de la plupart des maladies cardiaques. Le schéma d'extraction de caractéristiques proposé dans ce travail

Chapitre 3 : Traitement de signal ECG

utilise les amplitudes, les intervalles de temps, la largeur et les paramètres AR du signal ECG pour une analyse et une identification ultérieures. Aux fins de l'étude comparative, le système proposé dans ce travail suit la procédure d'extraction de dix-huit caractéristiques de chaque battement ECG répertorié dans le tableau 3.1.

Temporel	<p>PQ Intervalle de temps entre les ondes P et Q.</p> <p>PS Intervalle de temps entre les ondes P et S.</p> <p>PR Intervalle de temps entre les ondes P et R.</p> <p>TS Intervalle de temps entre les ondes T et S.</p>	<p>TQ Intervalle de temps entre les ondes T et Q.</p> <p>TR Intervalle de temps entre les ondes T et R.</p> <p>PT Intervalle de temps entre les ondes P et T.</p>
Amplitude	<p>P Hauteur d'amplitude de l'onde P.</p> <p>T Hauteur d'amplitude de l'onde T.</p> <p>S Hauteur d'amplitude de l'onde S.</p>	<p>R Hauteur d'amplitude de l'onde R.</p> <p>Q Hauteur d'amplitude de l'onde Q.</p>
Largeur	<p>QRS Largeur dans le temps de QRS.</p> <p>T Largeur dans le temps de l'onde T.</p>	<p>P Largeur dans le temps de l'onde P.</p>
Coefficients AR	<p>A1 Premier coefficient du modèle AR.</p> <p>A3 Troisième coefficient du modèle AR.</p>	<p>A2 Deuxième coefficient du modèle AR.</p>

Tableau 3.1 Liste des fonctionnalités extraites.

L'identification a été tentée en utilisant des caractéristiques analytiques extraites (amplitude, temps et largeur) et des caractéristiques de modélisation (paramètres AR). Les performances de la méthode proposée sont évaluées sur la base de la reconnaissance du sujet et de la fréquence de battement qui est déterminée par les sujets correctement classés de la base de données de test à l'aide d'algorithmes de classification «SVM».

3.3 Caractéristiques analytiques

Les caractéristiques analytiques sont généralement extraites en analysant et en localisant certains points de référence à partir des signaux ECG. Comme le pic R est déjà détecté en utilisant la détection QRS, les amplitudes Q, S, P et T sont localisées en trouvant des points locaux positifs et négatifs séparément autour des positions R de chaque fenêtre.

La **figure 3.2** Montre les résultats obtenus après le filtrage du signal ECG pour l'enregistrement 101 de la base de données MIT BIH. Il est clair qu'on a obtenu une bonne correction de la ligne de base et une meilleure élimination du bruit. Cependant ce filtre présente un inconvénient qui est un décalage de phase du signal résultant due au retard provoqué par le filtre.

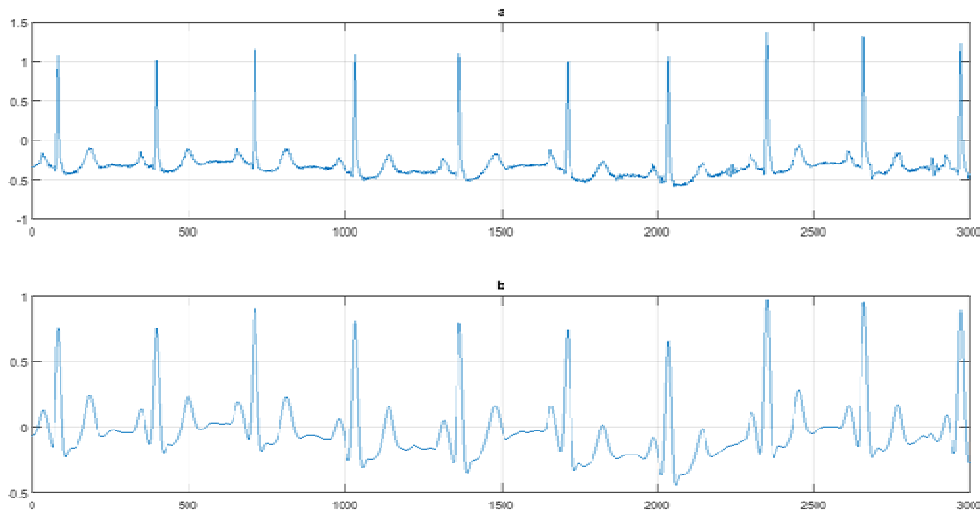


Figure 3.2 (a) signal ECG bruité, (b) signal ECG filtré (enregistrement 101)

La **figure 3.3** ci-dessous illustre les résultats obtenus pour la détection des pics R pour l'enregistrement 101 de la base de données MIT-BIH.

Chapitre 3 : Traitement de signal ECG

Le maximum positif à gauche du battement ECG est l'amplitude de l'onde P, et le négatif maximum à gauche est l'amplitude de l'onde Q. La même procédure est suivie pour trouver les amplitudes T et S sur le côté droit. Les figures 2.4a à 2.7 montrent les caractéristiques analytiques P, Q, S et T détectées et indiquées par un train d'impulsion au long du signal ECG.

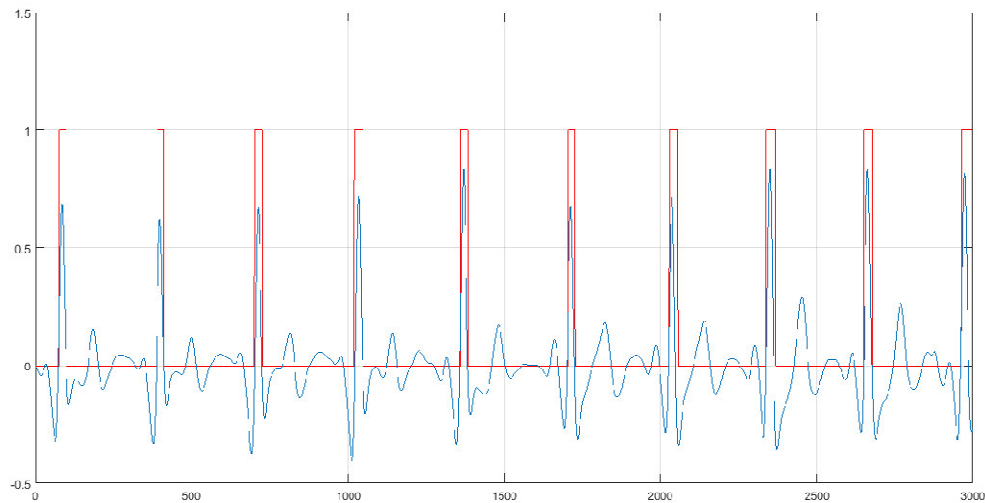


Figure 3.3 détection des pics R.

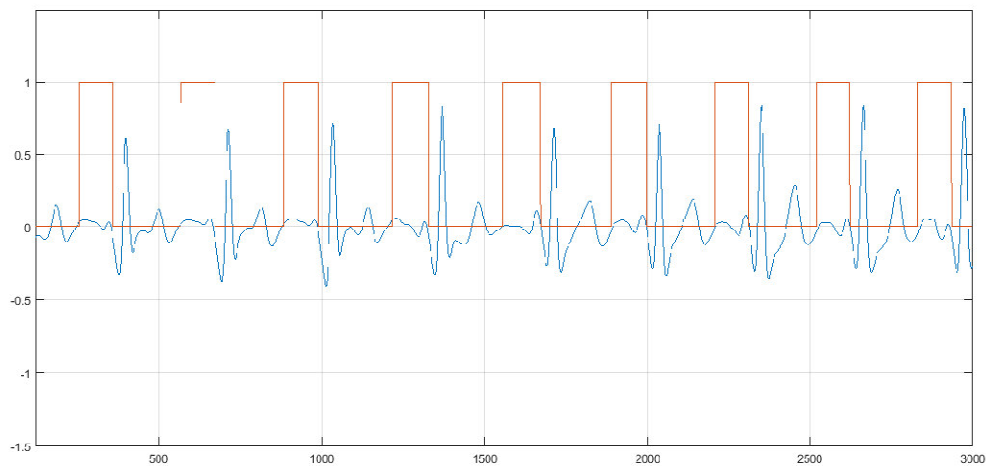


Figure 3.4 détection des ondes P

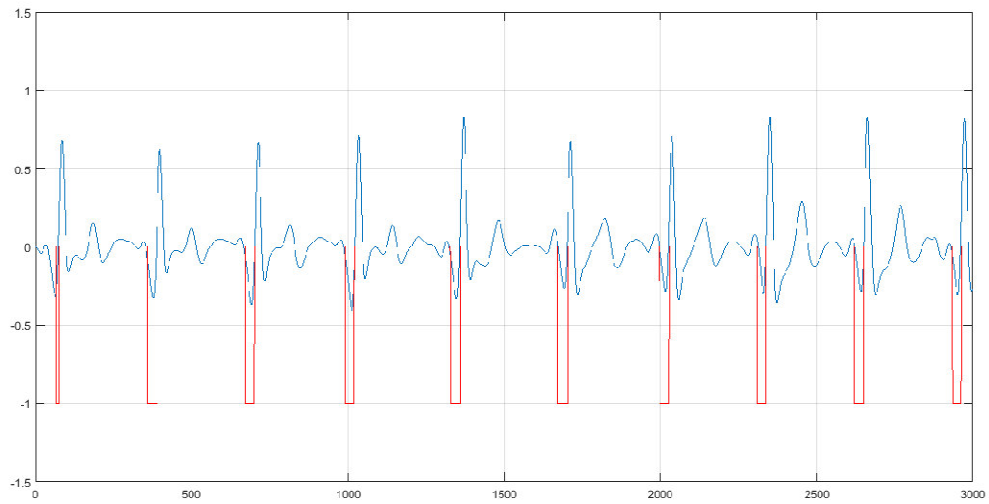


Figure 3.5 détection des ondes Q

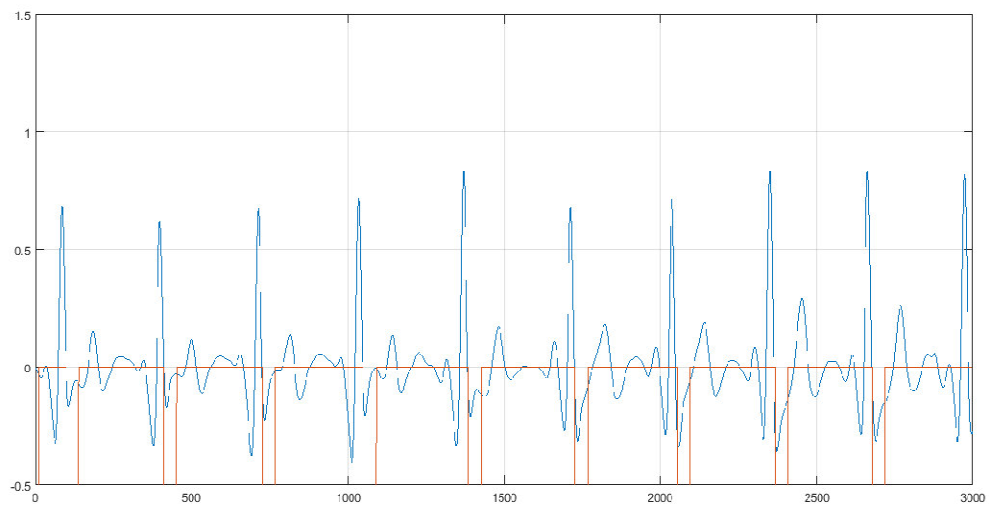


Figure 3.6 détection des ondes S

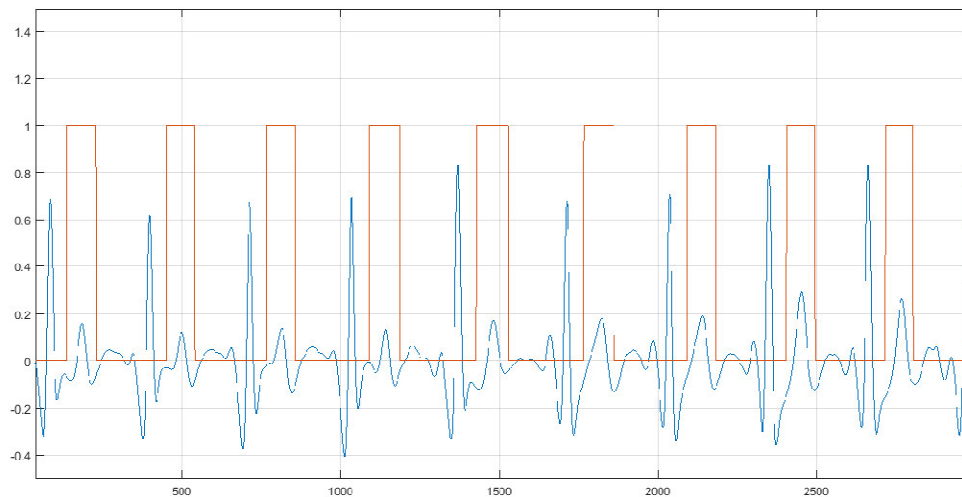


Figure 3.7 détection des ondes T

Dans ce travail, des distances temporelles, d'amplitude et de largeur entre les points indiqués ont été utilisées. Les paramètres extraits temporels et de largeur sont calculés à partir des points de référence indiqués sur la figure 3.3 à 3.7. La figure 3.8 montre la variabilité des différent intervalles tel que largeur de l'onde, intervalle PQ, intervalle QRS et intervalle QT par rapport intervalle RR.

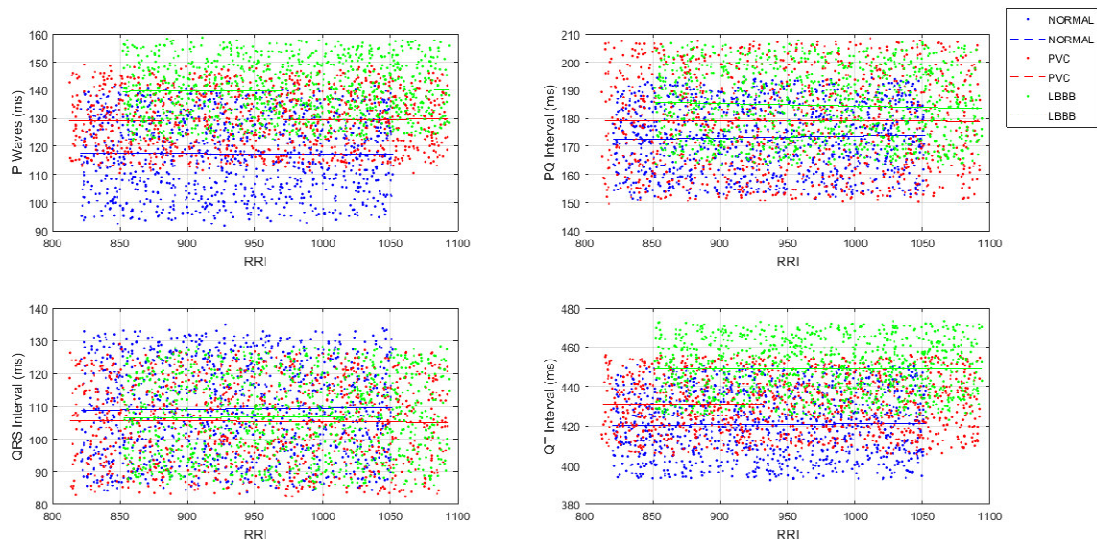


Figure 3.8 La variabilité des intervalles cardiaque par rapport RRI

3.4 Fonctionnalités de modélisation

En plus des fonctionnalités ECG analytiques décrites dans la section précédente, le système proposé utilise des fonctionnalités de modélisation obtenues par la modélisation de deux battements ECG successifs ou plus. En plus des fonctionnalités ECG analytiques décrites dans la section précédente, le système proposé utilise des fonctionnalités de modélisation obtenues par la modélisation de deux battements ECG successifs ou plus. Pour ce faire, une forme discrète d'un modèle de signal autorégressif (AR) d'ordre p , AR (p) est appliquée pour modéliser le signal ECG mesuré et les coefficients estimés ajoutés à un ensemble de fonctionnalités utilisé dans la dernière étape de classification de la proposition système. En utilisant le modèle AR proposé, une séquence de signaux échantillonnés $y(n)$ peut être représentée par la relation:

$$y(n) = a_1y(n-1) + a_2y(n-2) + \dots + a_py(n-p) + \mathcal{E}(n)$$

Où $a_k (k = 1, 2, \dots, p)$ sont les coefficients du modèle utilisés dans le processus de classification et $\mathcal{E}(n)$ est une série de bruit blanc, processus d'innovation avec une moyenne nulle et une variance σ^2 . Une modélisation précise de deux battements successifs à partir du signal ECG peut être obtenue en utilisant un modèle AR du 3ème ordre comme décrit plus haut dans ce paragraphe.

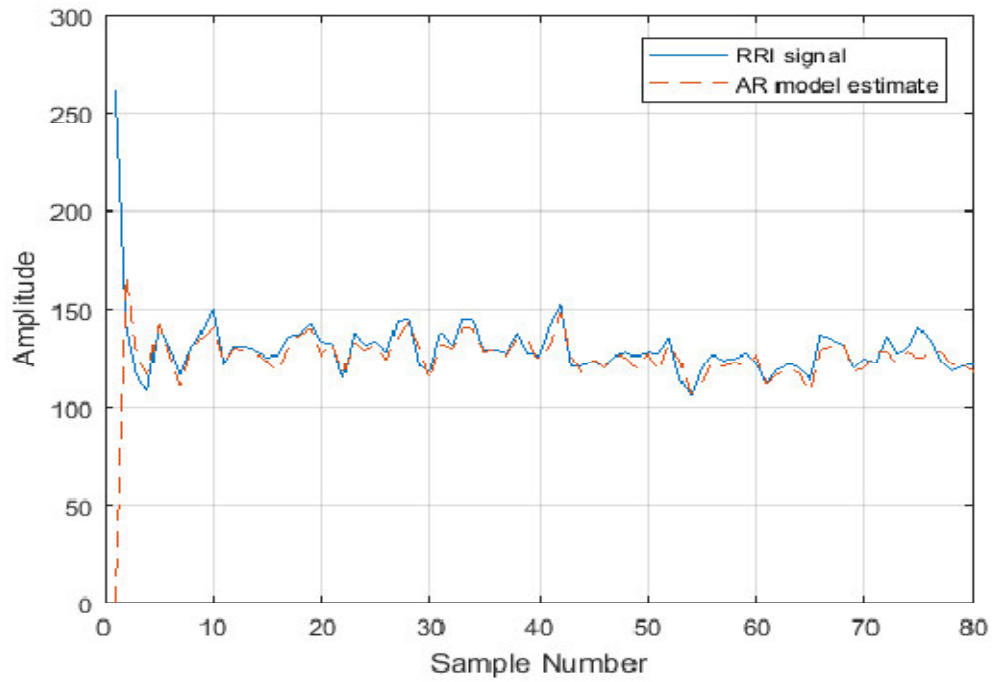


Figure 3.9 Le signal RR original (trait continu), et signal RR simule (trait discret) en utilisant le modèle AR.

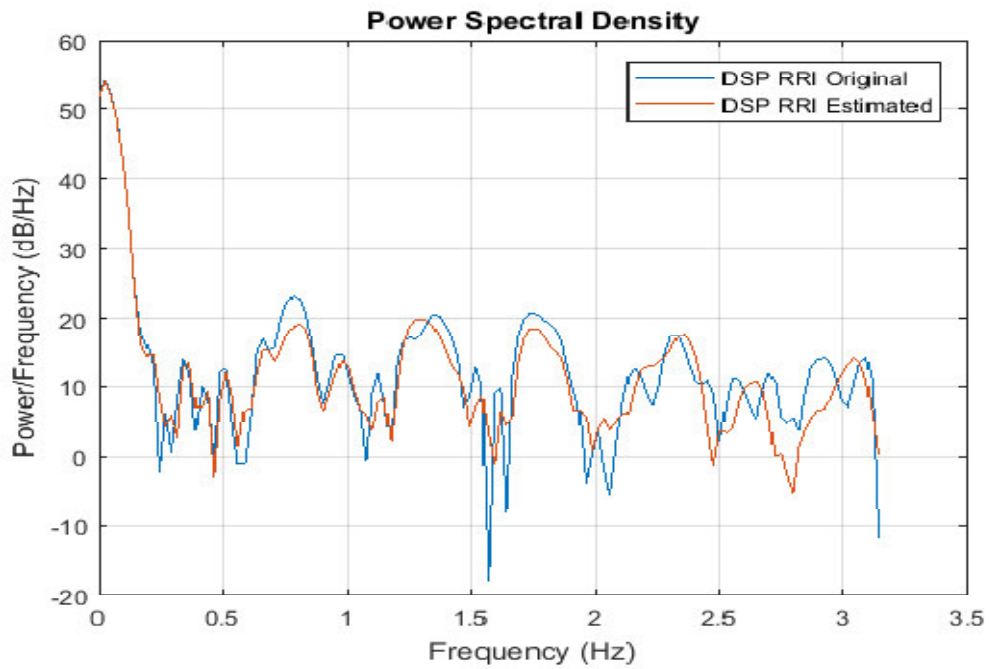


Figure 3.10 La DSP du signal RR original (trait continu) et signal RR simule (trait discret) en utilisant le modèle AR.

Un exemple d'application de ce modèle pour l'identification de la série RR correspondant au rythme cardiaque de l'enregistrement 101 est illustré dans la figure 3.9 alors que leur DSP estimée est présentée dans la figure 3.10.

L'algorithme développé a également été testé sur l'extension de la base de données MIT-HIB en utilisant des signaux ECG obtenus à partir des patients dans chacune des trois classes, y compris les signaux ECG normaux de type contractions ventriculaires prématurées (PVC) et blocs de branche gauche (LBBB). La figure 3.11 montre l'espace des caractéristiques 3D formé par 3 coefficients AR tel que a_1 , a_2 et a_3 extraits de cette base de données ECG.

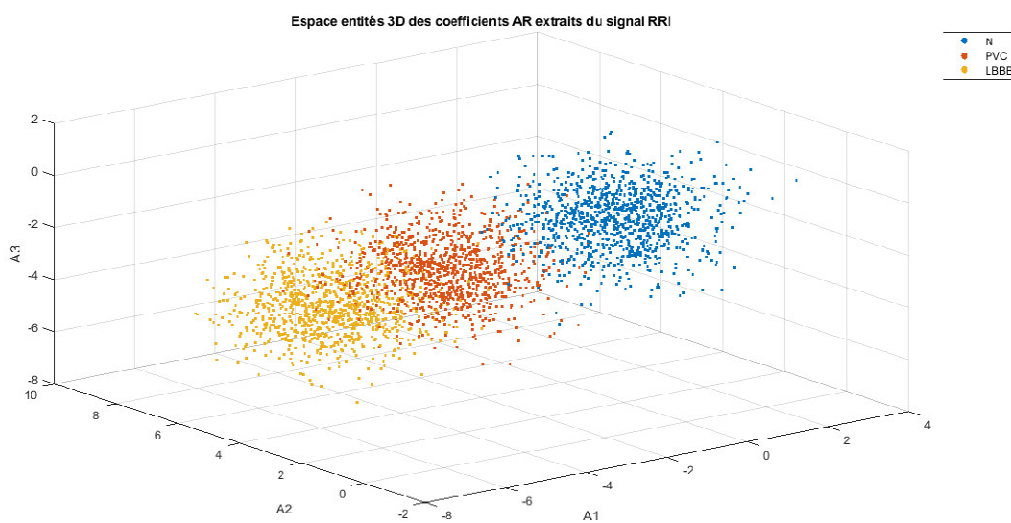


Figure 3. 11: Espace des caractéristiques 3D des coefficients AR extraits du signal RRI résultant en une classification complètement précise.

Conclusion :

L'examen de l'ECG a été largement utilisé pour diagnostiquer de nombreuses maladies cardiaques. Diverses techniques d'extraction de caractéristiques ont été proposées. Ce chapitre a donné un aperçu de la modélisation AR à utiliser comme techniques et algorithmes d'extraction de caractéristiques ECG. Cette approche peut atteindre une précision de 91% pour la modélisation du signal, une extraction rapide des caractéristiques en raison du faible nombre de paramètres extraits par rapport à d'autres méthodes d'extraction de caractéristiques et aboutit à une productivité de 99%. Les résultats de la modélisation AR pour représenter le signal RRI montrent jusqu'à 91% de précision de correspondance entre le signal autorégressif ECG et le signal ECG réel.

4.1. Introduction

Les SVMs reposent sur une théorie mathématique solide à l'inverse des méthodes de réseaux de neurones. Elles ont été développées au sens inverse du développement des réseaux de neurones : ces derniers ont suivi un chemin heuristique de l'application et l'expérimentation vers la théorie ; alors que les SVMs sont venues de la théorie du son vers l'application. Aujourd'hui, elles sont utilisées dans différents domaines de recherche et d'ingénierie tel que le diagnostic médical, le marketing, la biologie [34] [35] [36]. Dans ce chapitre, nous introduirons en première les notions sur l'apprentissage statistiques et le principe de machine à vecteur support, et on va introduisant cette classification et aussi interpréter les résultats obtenus.

4.2 Notion de l'apprentissage statistique

D'après Mari & Napoli (1996) [37], effectuer une classification, c'est mettre en évidence, d'une part, des relations entre des objets et, d'autre part les relations entre ces objets et leurs paramètres. Il s'agit de construire une partition de l'ensemble des objets en un ensemble de classes qui soient les plus homogènes possible. La classification, a donc deux objectifs à atteindre :

- Trouver un modèle capable de représenter la répartition des données (catégorisation).
- Définir de manière formelle l'appartenance à l'une ou l'autre des classes de toute nouvelle donnée (généralisation).

En effet la classification a pour but de réduire l'espace de recherche dans une base de données lors du processus d'identification.

4.3 Théorie des machines à vecteurs de support (SVM) :

Les Machines à Vecteurs de Support -SVM - ont été largement utilisées et appliquées aux problèmes de classification et la régression non linéaire. Nous nous intéressons ici à l'application des SVM dans le domaine bio-informatique et médical.

Les SVM ont été utilisées comme méthode de prédiction de cancer du sang (leucémie, lymphome, etc.) dans le contexte de la technologie des puces d'ADN (ou micro-array).

Liang et Lin ont prouvé l'efficacité des SVMs pour la classification des données ECG pour la détection d'une vidange gastrique tardive.

Une autre application en diagnostic médicale est présentée pour la détection des micro-calcifications dans les mammographies dans [38]

4.4 Principe des machines à vecteurs de support

Les SVM constituent une classe d'algorithmes basée sur le principe de minimisation de « Risque structurel » décrit par la théorie de l'apprentissage statistique qui utilise la séparation linéaire. Cela consiste à séparer par un hyperplan des individus représentés dans un espace de dimension égale au nombre de caractéristiques, les individus étant alors séparés en deux classes. Cela est possible quand les données à classer sont linéairement séparables. Dans le cas contraire, les données seront projetées sur un espace de plus grande dimension qu'elles deviennent linéairement séparables [39].

4.4.1 Cas de données linéairement séparables

Les cas de données linéairement séparables sont les plus simples des SVMs car ils permettent de trouver facilement le classificateur linéaire.

Dans la plupart des problèmes réels il n'y a pas de séparation linéaire possible entre les données, le classificateur de marge maximale ne peut pas être utilisé car il fonctionne seulement si les classes de données d'apprentissage sont linéairement séparables.

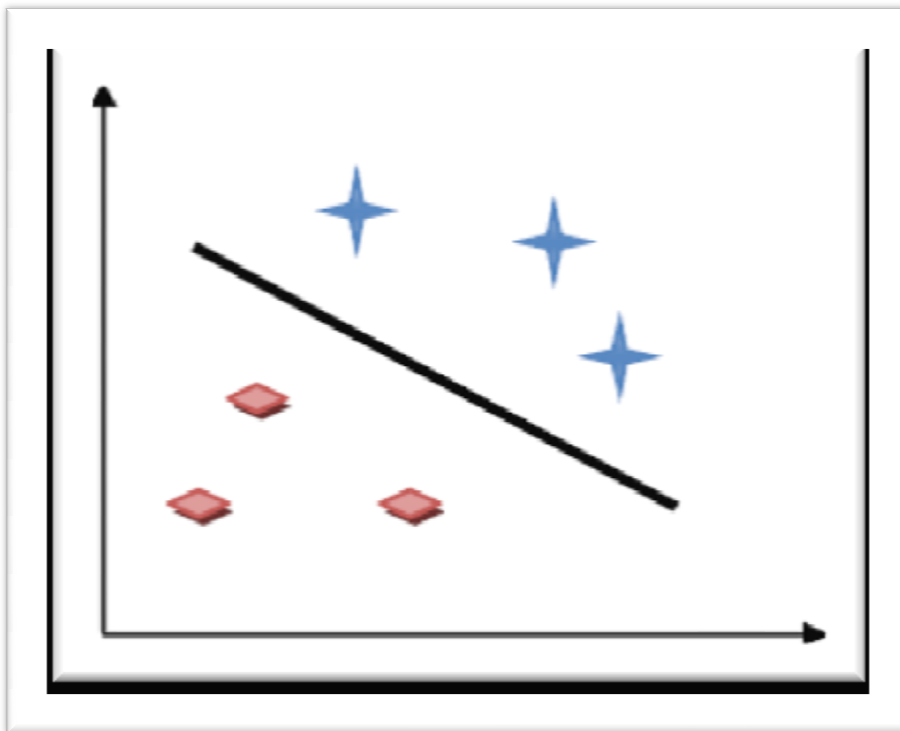


Figure 4.1 : Cas linéairement séparable.

4.4.2 Cas de données non-linéairement séparables

Pour surmonter les inconvénients des cas non linéairement séparables, l'idée des SVMs est de changer l'espace des données. La transformation non linéaire des données peut

permettre une séparation linéaire des exemples dans un nouvel espace. On va donc avoir un changement de dimension. Ce nouvel espace est appelé « espace de re-description ».

En effet, intuitivement, plus la dimension de l'espace de re-description n'est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les exemples est élevée.

On a donc une transformation d'un problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de re-description de plus grande dimension. Cette transformation non linéaire est réalisée via une fonction noyau.

En pratique, quelques familles de fonctions noyau paramétrables sont connues et il revient à l'utilisateur de SVM d'effectuer des tests pour déterminer celle qui convient le mieux pour son application. On peut citer les exemples de noyaux suivants : polynomiale, gaussien, sigmoïde et laplacien [40] [41].

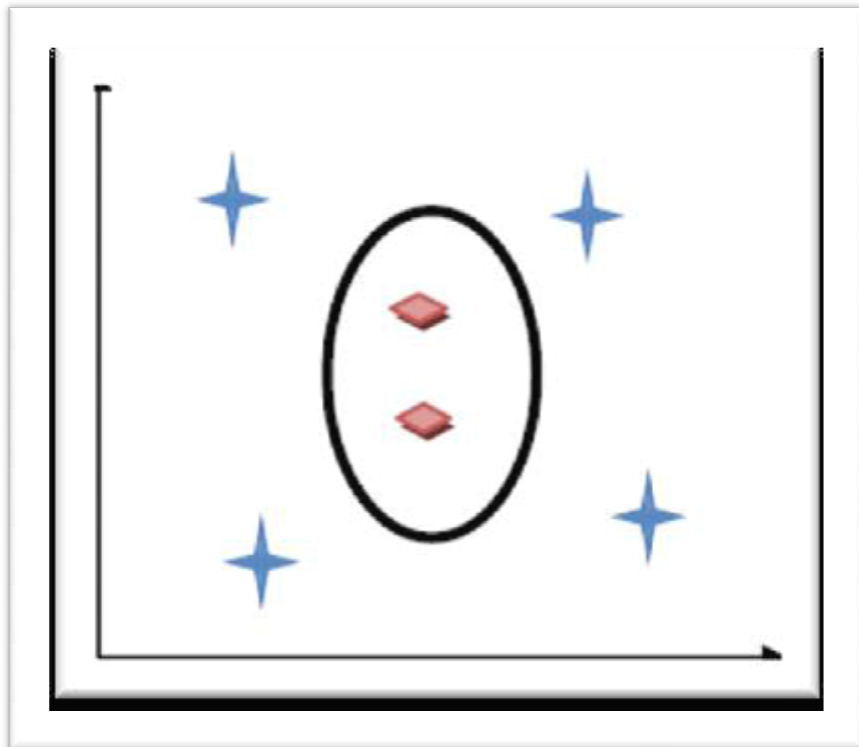


Figure 4.2 : cas non linéairement séparable.

4.5 La classification SVM

Les machines à support de vecteurs sont parmi les techniques les plus connues, développées par Vapnik en 1995, elles sont considérées comme une alternative récente pour la classification. Elles se basent principalement sur l'utilisation des fonctions appelées kernel, qui facilite la séparation des données [40].

En général, les SVM peuvent être utilisées pour résoudre plusieurs problèmes réels, tels que la classification textuelle et la régression. Pour cela, on doit construire une fonction f qui accepte un vecteur d'entrée x et qui retourne un vecteur de sortie y , avec:

$$y = f(x)$$

4.5.1 Formulation d'un problème de classification (supervisé)

Le classificateur doit estimer une fonction $f(x)$ qui est l'estimation de la fonction qui représente la relation entre l'objet et sa catégorie. Cette fonction est appelée fonction de décision :

$$F : X \rightarrow Y$$

X : L'ensemble des objets à classifier (appelé espace d'entrée).

Y : L'ensemble des catégories (appelé espace d'arrivée).

4.5.2. Minimisation du risque structurel

Deux types de données sont utilisés pour un problème d'apprentissage : les données d'entraînement (données d'origine pour calculer le modèle) et les données de test (pour évaluer la performance de généralisation du modèle). La qualité de ce modèle est alors jugée par rapport à sa capacité à réduire l'erreur de test ou de « généralisation ». Cependant, comme le modèle n'est pas construit en utilisant l'ensemble de test, l'erreur de généralisation ne peut pas être évaluée exactement car elle dépend de la distribution de probabilité des données. Suivant la théorie de Vapnik [42], nous supposons que les données sont générées selon une distribution de probabilité inconnue $P(x,y)$. De plus, nous supposons que les données sont indépendantes et identiquement distribuées (iid).

L'erreur moyenne commise sur toute la distribution $P(x,y)$ par la fonction $f(x)$ est donnée par:

$$R[f] = \int \frac{1}{2} Q(x) dp(x,y) \quad (4.1)$$

Où:

Q : est la fonction d'erreur (erreur absolue dans le cas des SVM).

x : est le vecteur d'entrée.

y : est l'ensemble des classes. Ainsi la fonction f devra être optimale : la fonction f_{opt} devra être calculée de sorte que l'erreur moyenne sur toute la distribution soit minimale.

$$F_{opt} = \operatorname{argmin}_f (R[F]) \quad (4.2)$$

Le critère formulé dans (2) est malheureusement inutilisable en pratique. En effet, pour calculer le risque, nous devrions disposer d'une estimation de la distribution $P(x,y)$, ce qui n'est pas le cas. La seule information dont nous disposons comme évaluation de l'erreur est l'erreur d'entraînement appelée Risque empirique :

$$Remp = \frac{1}{2} \sum_i^l \frac{1}{2} Q(x_i) \quad (4.3)$$

Où :

- l est le nombre d'objets d'entrainements.

Ceci n'est pas suffisant. La raison en est que l'on peut facilement trouver un modèle minimisant l'erreur d'entrainement mais pour lequel l'erreur de généralisation sera très grande. Donc cette dernière est liée à la famille de fonction utilisée comme modèle. Cette dépendance est nommée « risque structurel ». Dans la théorie de l'apprentissage statistique, Vapnik et Chervonenkis ont prouvé qu'il est possible de définir une majoration du risque structurel en fonction de la famille de fonctions utilisée pour le modèle [42]. L'une de ces majorations peut être calculée en utilisant la dimension de Vapnik–Chervonenkis (dimension VC) qui représente le plus grand nombre de points pouvant être séparés de toutes les façons possibles par un membre de l'ensemble de fonctions de F . Cela veut dire qu'il doit exister une configuration de $h = (VC(f))$ points, telle que les fonctions $f \in F$ peuvent leur assigner les 2^h combinaisons des labels (classes) possibles.

4.5.3 Architecture d'un classificateur SVM

4.5.3.1 La fonction noyau :

Afin de résoudre le problème de données non linéaires, la fonction noyau joue le rôle central de liaison des vecteurs d'entrées à l'espace de caractéristiques de grande dimension. (Voir figure 4.3).

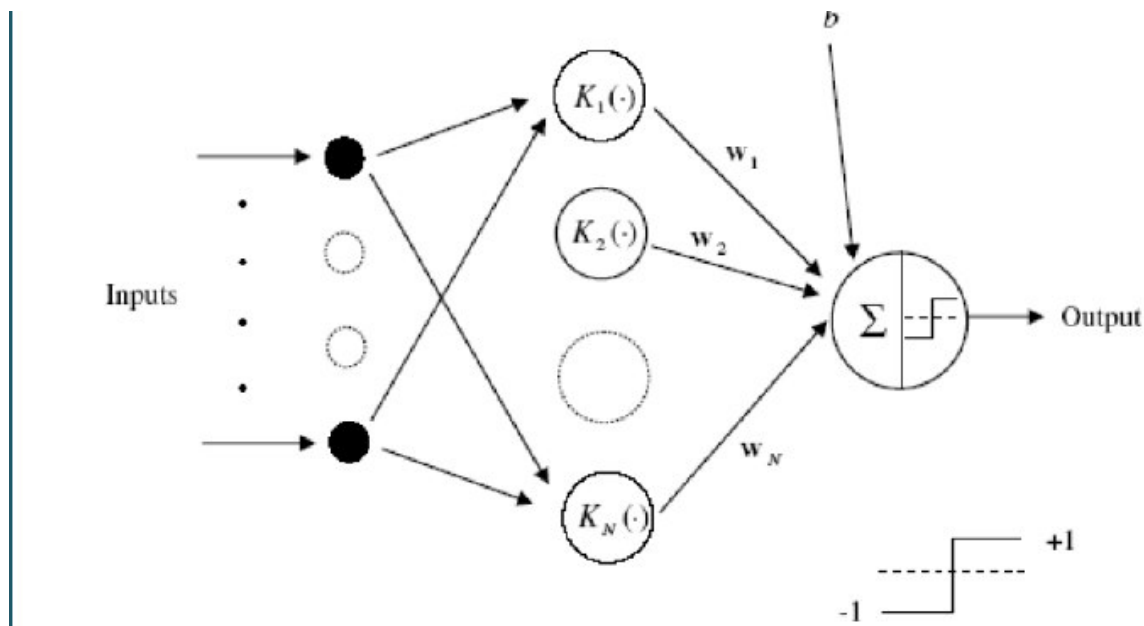


Figure 4.3 Architecture d'une machine à vecteurs de support (d'un nombre N)

Les choix typiques pour la fonction noyau sont :

Noyau gaussien à base radiale (ou RBF : Radial Basis Function) :

$$K(x, x_i) = \exp [-\gamma \|x - x_i\|] \quad (4.4)$$

Noyau polynômiale :

$$K(x_i - x_j) = (x_i^T x_j + 1)^p \quad (4.5)$$

Avec p : une constante qui spécifie le degré du polynôme.

4.5.3.2 Sélection de modèle SVM :

Une machine à vecteur de support binaire sépare les exemples positifs des exemples négatifs dans la phase d'apprentissage.

Les multiplicateurs de Lagrange α_i pour chaque machine binaire sont déterminés par

La minimisation de la fonction coût donnée par la formule (4.6) suivante :

$$P(w) = \frac{1}{2} \|w\|^2 \quad (4.6)$$

Ce problème est résolu à partir de la forme duale qui est exprimée par :

$$L_D = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j x_i x_j y_i y_j \quad (4.7)$$

Cela revient à chercher : $\sum_i \alpha_i y_i = 0$ et $0 \leq \alpha_i \leq C$

Si la valeur du paramètre de régularisation « C » qui contrôle la tolérance aux erreurs de classification dans la phase d'apprentissage est élevée, donc plus de pénalité sera donnée à l'erreur. Le vecteur d'apprentissage x_i qui a une valeur de α_i non nul est appelé « **vecteur de support** » [38].

4.5.3.3 Estimation de l'erreur de généralisation

La technique la plus populaire pour l'estimation de l'erreur de généralisation est la validation croisée (en anglais cross validation ou Leave One Out (LOO)) qui est utilisée indépendamment de la nature de la machine d'apprentissage utilisée. Le principe de cette méthode consiste à séparer en N sous-ensembles les échantillons de la base d'apprentissage, à apprendre sur $N - 1$ sous-ensembles, à valider sur le sous-ensemble restant, puis à faire tourner les sous-ensembles de façon à ce que chacun ait pu être testé. Au final, on note le nombre d'erreurs de classification de la procédure LOO par $l(X_1, Y_1, \dots, X_n, Y_n)$.

Il a été démontré dans [43] que cette procédure donne une estimation presque non biaisée de l'espérance de l'erreur de généralisation.

En effet, l'espérance $E(\cdot)$ de la probabilité p_{err}^{n-1} de l'erreur de test par une machine entraînée à partir de $n-1$ exemples est donnée par la formule (4.8):

$$E(p_{err}^{n-1}) = \frac{1}{n} E(l(X_1 Y_1 \dots X_n Y_n)) \quad (4.8)$$

Cependant la procédure est coûteuse en calcul, car nécessitant « n » apprentissages. Une procédure simple d'estimation d'erreur de généralisation est la validation croisée dite K -fold. Elle consiste à diviser l'ensemble des données en k sous-ensembles mutuellement exclusifs de taille approximativement égale.

L'apprentissage de la machine est effectué en utilisant $k-1$ sous ensemble et le test est effectué sur le sous-ensemble restant. Cette procédure est utilisée une fois pour le test. La moyenne des k taux d'erreur obtenus estime l'erreur de généralisation.

Nous avons montré dans notre travail que la procédure de la validation croisée donne de bonnes estimations de l'erreur de généralisation. Le nombre d'ensembles utilisé dans notre travail est égal à 5 (5fold) [38].

4.6. Les Machines à Vecteurs Supports (SVM) pour la régression

Lorsque les SVM sont utilisés dans des problèmes de régression pour prédire des valeurs réelles, on parle des SVR (Support Vector Regression). « Les SVM peuvent également être mis en œuvre en situation de régression, c'est-à-dire pour l'approximation de fonctions quand Y est quantitative. Dans le cas non linéaire, le principe consiste à rechercher une estimation de la fonction par sa décomposition sur une base fonctionnelle. La forme générale des fonctions calculées par les SVM se met sous la forme :

$$\phi(x, W) = \sum_{i=1}^{\infty} W_i V_i(x) \quad (4.9)$$

Le problème se pose toujours comme la minimisation d'une fonction coût, mais plutôt que d'être basée sur un critère d'erreur quadratique (moindres carrés), celle-ci s'inspire des travaux de Huber sur la recherche de modèles robustes et utilise des écarts absolus. On note ρ la fonction qui est paire, continue, identiquement nulle sur l'intervalle $[0, \epsilon]$ et qui croît linéairement sur $[\epsilon, +\infty]$. La fonction coût est alors définie par :

$$E(W, \gamma) = \frac{1}{n} \sum_{i=1}^n |Y_i - \phi(x_i, W)| + \gamma \|W\|^2 \quad (4.10)$$

Où γ est, comme en régression rigide, un paramètre de régularisation assurant le compromis entre généralisation et ajustement. De même que précédemment, on peut écrire les solutions du Problème d'optimisation. Pour plus de détails, se reporter à Schölkopf et Smola (2002). Les points de la base d'apprentissage associés à un coefficient non nul sont là encore nommés vecteurs support. Dans cette situation, les noyaux k utilisés sont ceux naturellement associés à

la définition de bases de fonctions. Noyaux de splines ou encore noyau de Dériclet associé à un développement en série de Fourier sont des grands classiques. Ils expriment les produits scalaires des fonctions de la base. » [44].

4.6.1. Formulation du problème dans le cas linéaire :

Nous allons à présent décrire les machines à vecteurs supports sous leur forme conventionnelle. On dispose d'un ensemble de données $D = \{(X_i, Y_i) \in X \times Y\}_{i=1}^N$, où $X \subset \mathbb{R}^d$ et $Y \subset \mathbb{R}$. L'objectif est de trouver une fonction f dans un espace de Hilbert H , par minimisation du risque structurel (SRM), établissant une relation entre les variables X et la grandeur à modéliser y , à partir de l'ensemble de mesures D . La fonctionnelle que l'on cherche à optimiser s'écrit :

$$F(f) = c \sum_{i=1}^N |Y_i - F(x_i)|_\varepsilon + \frac{1}{2} \|F\|_H^2 \quad (4.11)$$

Où $\|\cdot\|_H$ est une norme sur H , et où $|\cdot|_\varepsilon$ Représente la fonction de coût ε -insensible suivante

$$V(Y, f(x)) = |Y - f(x)|_\varepsilon = \begin{cases} 0 & \text{si } |x| < \varepsilon \\ |x| - 1 & \text{sinon} \end{cases}$$

Cette fonction de coût ne pénalise que les déviations supérieures à ε entre les sorties mesurées et estimées. Nous allons, dans un premier temps, réaliser un modèle linéaire, en cherchant la fonction de régression f sous la forme suivante : $f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b$, où $\mathbf{w} \cdot \mathbf{x}$ est un produit scalaire dans l'espace X [39].

4.6.2. Extension au cas non linéaire

Il est possible de traiter le problème non linéaire de manière essentiellement identique à ce qui vient d'être fait dans le cas linéaire, en cherchant la solution dans un RKHS muni d'un noyau convenablement choisi. Ceci revient à effectuer un « prétraitement » des variables, qui consiste en une transformation non linéaire de l'espace des variables X vers un espace image F , grâce à une fonction $\varphi : X \rightarrow F$ associée au noyau K . Le noyau choisi définissant la « distance » entre les objets dans l'espace des caractéristiques.

4.6.3. Choix des hyper paramètres

Les différents paramètres relatifs au noyau utilisé (écart type des gaussiennes dans le cas de noyaux gaussiens) Afin de choisir au mieux ces hyper paramètres, des méthodes classiques peuvent être mises en oeuvre (notamment la validation croisée). Il est également possible, dans ce cas particulier des SVM, d'établir des bornes à certains indices de performance (erreur de validation, erreur de leave-one-out, ...), et d'optimiser ceux-ci par une descente de gradient vis-à-vis des hyper paramètres [45].

Chapitre 4 : Classification SVM

La solution obtenue par un algorithme de SVM, c'est-à-dire l'ensemble des coefficients $\{\alpha_i, \alpha_i^*\}_{i=1}^N$, dépend évidemment des hyperparamètres utilisés [45].

4.7 Résultats

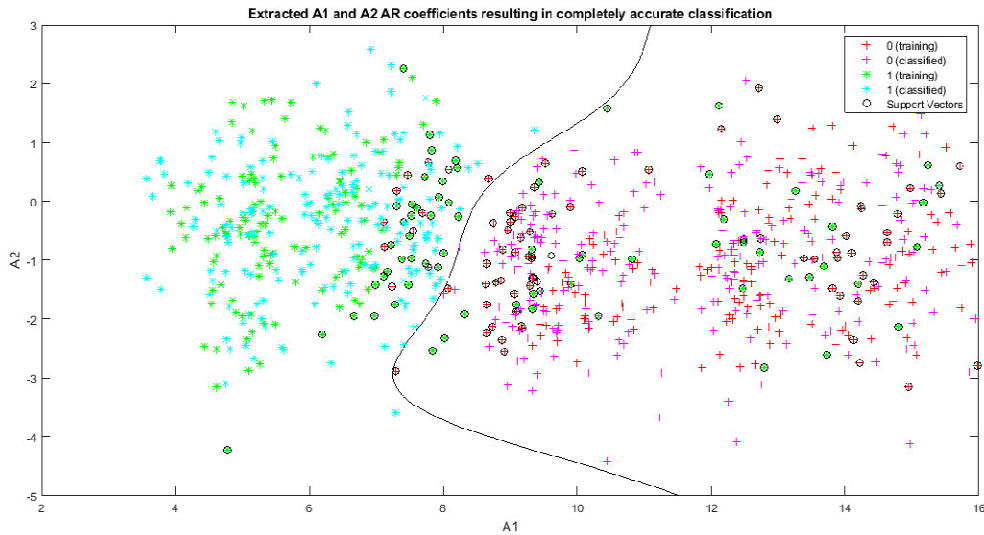


Figure 4.4 Coefficients A1 et A2 AR extraits résultant en une classification complètement précise à l'aide de SVM

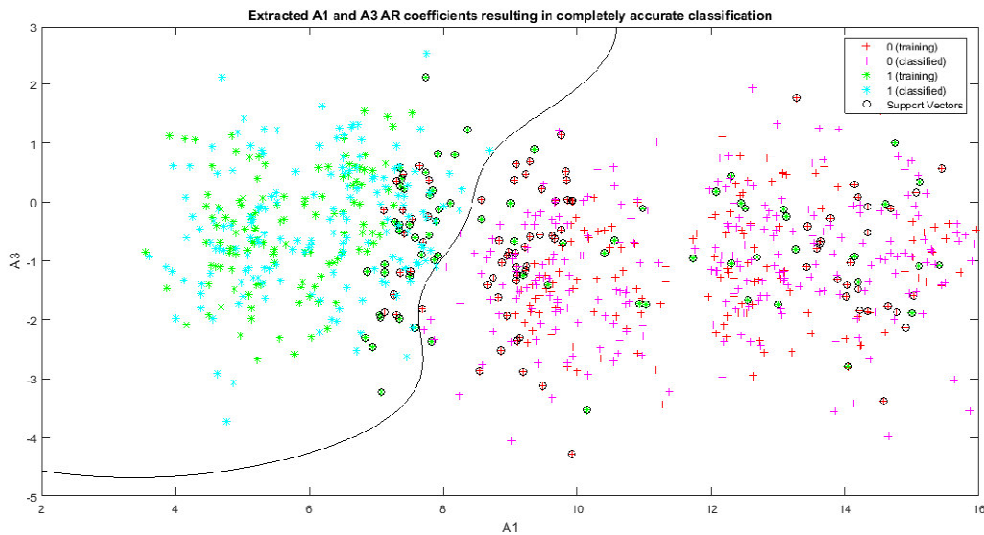


Figure 4.5 Coefficients A1 et A3 AR extraits résultant en une classification complètement précise à l'aide de SVM

Chapitre 4 : Classification SVM

Type de classification (SVM)	Précision	Valeur productive négative (NP)
Onde P	55.56 %	33 %
Intervalle QRS	78.89 %	56 %
Intervalle PQ et QT	91.89 %	21%
Coefficients AR (A1, A2 et A3)	95 %	17.84 %

Tableau 4. 1 La précision (%) obtenu dans la classification correspondant à l'algorithme biométrique

La Figure 4.4, la Figure 4.5 et le tableau 4. 1 montrent les résultats de classification de SVM équipés du coefficient AR de la normale en vert et de deux arythmies en rouge. Les résultats graphiques de la classification illustrent une précision de 95% en utilisant SVM.

Conclusion

Dans ce chapitre, une classification est réalisée en utilisant la technique machines à vecteurs de support (SVM).

La classification d'électrocardiogramme proposée et les systèmes biométriques de ce travail ont atteint un taux de classification correct de 95% en identifiant les signaux d'électrocardiogramme normaux et anormaux et chaque personne individuellement à partir de son signal d'électrocardiogramme.

Références bibliographiques

- [1] N. BAKEK YETTOU, «modélisation paramétrique et non paramétrique en vue de l'identification de système cardiaque », thèse de doctorat, université ABOU BEKR BELKAID, Tlemcen, 2015.
- [2] Z. Bouabida. « Analyse et Traitement du Signal ECG par la Décomposition Modale Empirique ». Mémoire de magister en électronique biomédicale Université ABOU BEKR BELKAID, Tlemcen, 2009.
- [3] A. Cabasson. « Estimation et analyse des intervalles cardiaques. Traitement du signal et de l'image [eess.SP »]. Université Nice Sophia Antipolis, 2008. Français.
- [4] Support de Cours (Version PDF) - Collège National des Enseignants de Cardiologie et Maladies Vasculaires. Université Médicale Virtuelle Francophone.2010-2011.
- [5] https://www.neurones.espci.fr/Theses_PS/DUBOIS_R/Chapitre1.pdf
- [6] M. Oltean, Rapport de stage, Traitement automatique de signaux ECG, Ecole Nationale Supérieure de Télécommunications de Bretagne, 2006.
- [7] M. MESSIOUD, « classification des signaux ECG en utilisant les réseaux de neurones », Mémoire de fin d'étude, université 8 Mai 1945, Guelma, 2019.
- [8]https://www.researchgate.net/profile/Benali_Radhwane2/publication/281237494/figure/fig/2/AS:363025178611714@1463563356019/Montage-de-Goldberger-pour-lenregistrement-des-derivations-unipolaires-des-membres.png.
- [9] (<https://www.my-ekg.com/imag/index-ico/generalidades/ekg-leads.png>)
- [10] K. BENSAFIA. « transmission sans fil par voie GSM et traitement de signal électrocardiographique (ECG) », mémoire de magister, université Mouloud MEMMERI, Tizi-Ouzou.
- [11] R. BENALI. « Analyse du signal ECG par réseau adaptatif d'ondelettes en vue de la reconnaissance de pathologies cardiaques », thèse de doctorat en Electronique Biomédicale, université ABOU BEKR BELKAID, Tlemcen, 2013.
- [12] M. TALBI. « Traitement de signal électrocardiographique (ECG) ». Thèse de doctorat en sciences. Université Mentouri, Constantine, 2011.

Références bibliographiques

- [13] M. Nait-Hamoud, « segmentation et classification de signal ECG par les SVM Flous Multi-classes ». Mémoire de magister .Université Cheikh Larbi Tbessi, Tébessa, 2010.
- [14] J. Adamec, R. Adamec, «ECG holter: manuel d'interprétation electrocardiographique», Edition Médecine et Hygiène, 2000.
- [15] A. Belgacem, « Classification des signaux EGC avec un système-multi-agent neuronale ». Mémoire de magister. ABOU BEKR BELKAID, Tlemcen, 2012.
- [16] MIT-BIH Arrhythmia Database, <http://www.physionet.org/physiobank/database>.
- [17] A. Ouali, « Contributions à l'analyse et au traitement des signaux biomédicaux non stationnaires Applications aux électrocardiogrammes et électroencéphalogrammes », Thèse de doctorat en Informatique et traitement de signal, Université Sultan Moulay Slimane, MAROC ,2015.
- [18] S. Canan, Y. Özbay and B. Karlik, "A method for removing low varying trend from ECG signal," Proceedings of the International Conference in Biomedical Engineering Days, pp 144–146, 1998.
- [19] M.L. Riccio and J.C. Belina, "A versatile design method of fast, linear-phase FIR filtering systems for electrocardiogram acquisition and analysis systems," Computers in Cardiology, pp 147–150, 1992.
- [20] J.M. Leski and N. Henzel, "ECG baseline wander and powerline interference reduction using nonlinera filter bank," Signal Processing, 85, pp 781–793, 2005.
- [21] L. Sörnmo and P. Laguna, "Electrocardiogram ECG signal processing," Wiley Encyclopedia of Biomedical Engineering, 2006.
- [22] L. Sörnmo, "Time-varying digital filtering of ECG baseline wander," Medical and Biological Engineering and Computing, 31, pp 503–508, 1993.
- [23] L. Sörnmo, "Time-varying filtering for removal of baseline wander in exercise ECGs," Computers in Cardiology, pp 145–148, 1991.
- [24] P.S. Hamilton, "A comparison of adaptative and nonadaptative filters for reduction of power line interference in the ECG," IEEE Trans. Biomed. Eng., 43, pp 105–109, 1996.

Références bibliographiques

- [25] Jiapu Pan and Willis J. Tompkins, A Real-Time QRS Detection Algorithm, IEEE Transactions on Biomedical Engineering, Vol. BME- 32, NO. 3, 1985.
- [26] Z.E. Hadj Slimane and F. Bereksi Reguig, Detection of the QRS complex by linear prediction. Journal of Medical Engineering & Technology, Vol. 30, NO.3, PP 134-138, 2006.].
- [27] P Laguna, N. V. Thakor, P. Caminal, R Jane, Hyung-Ro Yoon, A Bayes de Luna, V.Marti, JosepGuindo, new algorithm for QT interval analysis in 24 hour holter ECG:performance and applications. Med & Biol Eng & Comput. vol 28, pp 67-73. 1990.
- [28] L.A. Dobrzanski , R. Honysz , S. Fassois , On the identification of composite beam dynamics based upon experimental data, Journal of Achievements in Materials and Manufacturing Engineering. Vol 16, Issue 1-2 .May-June- 2006.
- [29] H. Akaike, A new look at the statistical model identification, IEEE Trans. Autom. Contr, 19, pp. 716-723. (1974).
- [30] A. Daniel. Finan, Francis J. Doyle, Cesar C. Palerm, Wendy C. Bevier, Howard C. Zisser, Lois Jovanovič, and Dale E. Seborg, “Experimental Evaluation of a Recursive Model Identification Technique for Type 1 Diabetes” J Diabetes Sci Technol. 2009 September; 3(5): 1192–1202
- [31] Kay, S. M. (1988):”Modern spectral analysis: theory and application” (prentice Hall, Englewood Cliffs, New Jersey).
- [32] A. Porta, G. Baselli, E. Caiani, A. Malliani, F. Lombardi, and S. Cerutti, “Quantifying electrocardiogram RT-RR variability interactions,” Med. Biol. Eng. Comput., vol. 36, no. 1, pp. 27–34, 1998.
- [33] N.Y. Baakek, F. Bereksi Reguig, Z. E. Hadj Slimane, Dynamic linear parametric models for the heart rate signal identification, Biomedical engineering international conference (BIOMEIC'12) — October 10-11, 2012, Tlemcen (Algeria).
- [34] : I.M. GUYUN et B. BOSER et V. VAPNIK. « Automatic capacity tuning of very large vc-dimension classifiers ». In Advances in neural information processing systems, 1993.p147-155.
- [35] : B.E. BOSER, I.M. GUYUN, et V.N. VAPNIC. « A training algorithm for optimal

Références bibliographiques

margin classifiers». In Proceedings of the fifth annual workshop on Computational learning theory, ACM, 1992, p 144–152.

[36] : A. DJEFFAL. «Utilisation des méthodes Support Vector Machine (SVM) dans l'analyse des bases de données ». Thèse de doctorat, université de Biskra(Algérie) ,2012.

[37] A. Tezmoz, H. Sari-Sarraf, S. Mitra, R. Long. "Customized Hough Transform for Robust Segmentation of the Cervical Vertebrae from X-Ray Images". *SSIAI2002*, Proc. IEEE 1537, pp.224-228, 2002.

[38] S. Benyahia. « Application de la classe des méthodes d'apprentissage statistique SVM (support vector machine) pour la reconnaissance des formes dans les images », Mémoire de magister en Informatique, Université de Tlemcen, 2011.

[39] M. salem, « Prédiction de la Fréquence de Résonance d'une Antenne Microbande Avec une Ouverture dans le Plan de Masse », mémoire de master, UNIVERSITE MOHAMED BOUDIAF - M'SILA, 2016.

[40] H. MOHAMEDALLY et B. FORMANI : « SVM : Machines à vecteurs de support ou séparateurs à vastes marges ». BD Web, ISTY3, Versailles St Quentin. (2006).

[41] M. SENOUSSAOUI : « Application des modèles de Markov cachés & les machines à vecteurs de support pour la reconnaissance des caractères isolés d'écriture en ligne ».diplôme de magister, SIMPA, 2007.

[42] M.Benjamin, HOWE, B.S., B.S.E.E.,” Segmentation of cervical and lumbar vertebrae in X-ray images using active appearance models and extensions », Thèse de Master,Université Texas tech,2003.

[43] Cortes C, Vapnik VN; support vector networks Machine Learning, 1995

[44] <http://wikistat.fr/> ,20 mai 2016

[45] O. Chapelle, V.N. Vapnik, O. Bousquet, S. Mukherjee. « Choosing Multiple Parameters for Support Vector Machines ». Springer, 2002.

Conclusion générale

Conclusion générale

Ce travail propose et étudie le système de traitement et d'analyse ECG avec deux applications principales - le système de classification ECG capable d'identifier les signaux ECG normaux (sains) et anormaux (malsains) et le système biométrique ECG capable d'identifier personnes individuelles à partir du signal ECG mesuré. La détection du complexe QRS, des ondes T et P et une analyse automatique du rythme pour trouver les intervalles de temps et les amplitudes des différents battements du flux de signal ECG. Ce travail a également utilisé et fourni une analyse détaillée des techniques de modélisation du signal basées sur AR et leur application pour l'extraction des caractéristiques ECG de manière nouvelle et efficace. La méthode autorégressive garantit un moyen rapide et précis d'extraire les caractéristiques de signal pertinentes, tout en produisant un haut niveau de productivité. Une évaluation comparative des performances de différents algorithmes pour l'extraction de caractéristiques en comparant diverses techniques aux méthodes d'extraction AR utilisées dans ce travail. La précision de la modélisation de méthode AR dépend de l'ordre des modèles utilisés. La technique l'AIC a été examinée et utilisée dans ce travail pour déterminer et sélectionner les ordres optimaux de modèle AR dans la modélisation précise du signal ECG. Le modèle AR peuvent représenter des signaux ECG avec une précision allant jusqu'à 91%, c'est-à-dire une correspondance entre le signal ECG modélisé et réel en sélectionnant l'ordre de modèle approprié. Différents algorithmes de classification ont été testés pour classer les caractéristiques d'analyse et de modélisation extraites du signal ECG. Les performances élevées de la classification dépendent de la capacité à séparer les vecteurs d'entités dans l'espace d'entités. En raison de sa simplicité et de son efficacité, le classifieur SVM est l'une des méthodes les plus couramment utilisées en bioinformatique et dans d'autres domaines, mais il faut prendre soin de sélectionner l'ordre du modèle ainsi que les différentes métriques de distance. Les fonctionnalités d'analyse et de modélisation extraites du signal ECG sont intégrées dans les algorithmes de classification SVM pour effectuer une tâche de reconnaissance biométrique ECG avec un taux de classification correct à 95% ou une précision lorsque l'approche est testée sur divers signaux disponibles dans les bases de données ECG. Le modèle AR autorégressif ont été mis en œuvre afin d'effectuer des tâches de classification ECG et de systèmes biométriques avec une précision et une productivité supérieures par rapport aux autres techniques et méthodes suggérées dans la littérature.

Résumé

L'électrocardiogramme (ECG) est l'onde P-QRS-T représentant les informations sur l'état du cœur. La forme et la taille du signal ECG peuvent contenir des informations utiles sur la nature de la maladie affectant le cœur. Les médecins utilisent des électrocardiogrammes pour détecter et diagnostiquer des conditions telles que les arythmies et les infarctus du myocarde.

Le travail décrit dans ce mémoire étudie le système conçu pour deux applications principales, le système de classification d'électrocardiogramme basé sur des modèles autorégressifs qui identifie les signaux d'électrocardiogramme normaux et anormaux et le système biométrique d'électrocardiogramme basé sur des fonctionnalités d'analyse et de modélisation qui identifie chaque personne individuellement à partir de son électrocardiogramme.

Le but de toute classification d'électrocardiogramme et système biométrique décrits dans ce travail est d'obtenir un taux de précision élevé lors de l'identification des électrocardiogrammes. La classification de l'électrocardiogramme et les systèmes biométriques se composent de quatre étapes principales, le prétraitement du signal de l'électrocardiogramme, la détection complexe QRS, l'onde P et T, l'extraction de caractéristiques et les algorithmes de classification

Mots clés : ECG, l'onde P-QRQ-T, classification, système biométrique, modèle autorégressif, modélisation.

Abstract

The electrocardiogram (ECG) is the P-QRS-T wave representing information about the condition of the heart. The shape and size of the ECG signal can contain useful information about the nature of the disease affecting the heart. Doctors use electrocardiograms to detect and diagnose conditions such as arrhythmias and myocardial infarctions. The work described in this dissertation studies the system designed for two main applications, the ECG classification system based on autoregressive models that identifies normal and abnormal ECG signals and the ECG biometric system based on autoregressive models. Analysis and modeling that identifies each person individually from their electrocardiogram.

The goal of any ECG classification and biometric system described in this work is to achieve a high rate of accuracy when identifying ECGs. ECG classification and biometric systems consist of four main steps, ECG signal preprocessing, QRS complex detection, P and T wave, feature extraction and classification algorithms.

Keywords: ECG, P-QRQ-T wave, classification, biometric system, autoregressive model, modeling.