SAKARYA UNIVERSITY
INSTITUTE OF SCIENCE AND TECHNOLOGY

ANALYSIS AND DENOISING OF ECG SIGNALS

M.Sc. THESIS

Tasnim AHMED ABDELRAZIG MOHAMMED

Department	**:**	**ELECTRICAL AND ELECTRONICS ENGINEERING**
Field of Science	**:**	**ELECTRONICS**
Supervisor	**:**	**Professor Abdullah FERİKOĞLU**

March 2019

SAKARYA UNIVERSITY
INSTITUTE OF SCIENCE AND TECHNOLOGY

ANALYSIS AND DENOISING OF ECG SIGNALS

M.Sc. THESIS

Tasnim AHMED ABDELRAZIG MOHAMMED

Department	:	**ELECTRICAL AND ELECTRONICS ENGINEERING**
Field of Science	:	**ELECTRONICS**
Supervisor	:	**Professor Abdullah FERİKOĞLU**

This thesis has been accepted unanimously / with majority of votes by the examination committee on 18.03.2019.

Professor Abdullah FERİKOĞLU	**Assoc. Prof. Şuayb Çağrı YENER**	**Assist. Prof. Halil ARSLAN**
Head of Jury	**Jury Member**	**Jury Member**

DECLERATION

I declare that all the data in this thesis was obtained by myself in academic rules, all visual and written information and results were presented in accordance with academic and ethical rules, there is no distortion in the presented data, in case of utilizing other people's works they were refereed properly to scientific norms, the data presented in this thesis has not been used in any other thesis in this university or in any other university.

Tasnim AHMED ABDELRAZIG MOHAMMED

14.02.2019

ACKNOWLEDGMENTS

In The Name of Allah Most Gracious Most Merciful.First, I give thanks to God for protection and ability to do work. I am so grateful to the Turkey scholarship program for making it possible for me to study here. I give deep thanks to the Professors and lecturers at the faculty of Electrical and Electronics Engineering and other workers of Sakarya university. My special and heartily thanks to my supervisor, Professor Abdullah FERİKOGLU who encouraged and directed me. With his supervision that this work came into existence. For any faults I take full responsibility.

I also thank my family who encouraged me and prayed for me throughout the time of my research. I would like to thank my parents, whose love and guidance are with me in whatever I pursue. They are the ultimate role models. Most importantly, I wish to thank my loving and supportive husband, Ayman, and my wonderful baby, Elfatih who provide unending inspiration.

May the Almighty God richly bless all of you.

TABLE OF CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

ECG	: Electrocardiogram
BSS	: Blind source separation
ICA	: Independent component analysis
EKF	: Extended Kalman Filter
SA	: Sinoatrial node
AV	: Atrioventricular node
PLI	: Power line interference
BW	: Baseline wander
EMG	: Electromyogram
MA	: Muscle artifact
EM	: Electrode motion artifact
FASTICA	: Fast independent component analysis
FIR	: Finite impulse response
MIT-BIH	: Massachusetts institute of technology - beth israel hospital
tanh	: Hyperbolic tangent
RMSE	: Root mean square error
SNR	: Signal to noise ratio

LIST OF FIGURES

LIST OF TABLES

SUMMARY

Keywords: Electrocardiogram, Blind Source Separation, Independent Component Analysis, Extended Kalman Filter.

The Electrocardiogram (ECG) is a bioelectric signal generated from the contraction of the human heart muscle, which is recorded from the body surface using electrodes for the diagnosis purpose. During the acquisition or transmission, noises generated from the surrounded electrical equipment, the motion of the patient, the movement of the electrodes, or the contraction of the muscle around the heart usually interfere with the clinically recorded ECG signals. The interference of the noises with the recorded ECG signal in the spatial, temporal, and frequency domain mask the desired signal and reduce the amount of information which can be extracted from it, mislead the clinician, and obstruct the diagnosis process. Many filtering techniques have been proposed for reducing, extracting or separating the noise sources but finding a robust and reliable filtering technique, which completely remove the noise without distorting the morphological shape of the signal still form a challenge. Recently, Blind Source Separation (BSS) and model-based filtering techniques have shown good results in the area of biomedical signal processing. Here in this study, Robust Independent Component Analysis (RobustICA) has been used to separate the measured ECG signal mixture into independent components based on its statistical features then a clean signal has been estimated. On the other hand, model-based filtering has been done using the nonlinear dynamical model of ECG signal with the Extended Kalman Filter (EKF) to estimate a noise-free signal. The proposed algorithms have been used to remove muscle contraction artifact, baseline shift, and electrode motion artifact, which are the most common ECG noises. The efficiency of the proposed methods has been measured by using several noisy ECG signals generated by artificially adding baseline wander, muscle artifact, and electrode movement noises to normal ECG signal. This generates a database with a range of signal to noise ratio (SNR) from 20 to -20 Decibel to inspect the filtered ECG recordings by studying their SNR and morphology after the filtering process. Performance comparison of ICA and EKF are showed that ICA produces better results in reducing muscle artifact when compared with baseline wander and electrode movement artifacts reduction ratio, while EKF shows good results in the reduction of muscle artifact only when the SNR is low. Analysis results demonstrate that ICA is better than EKF in the reduction of baseline wander and electrode movement artifacts while both show good result in the reduction of muscle artifact.

EKG İŞARETLERİNİN ANALİZİ VE GÜRÜLTÜDEN ARINDIRILMASI

ÖZET

Anahtar kelimeler: Elektrokardiyogram, Kör Kaynak Ayırma, Bağımsız Bileşen Analizi, Genişletilmiş Kalman Filtresi,

Elektrokardiyogram (EKG), tanı amaçlı elektrotlar kullanarak vücut yüzeyinden kayıt yapılan, insan kalp kasının kasılmasından kaynaklanan bir biyoelektrik sinyaldir. Elde etme veya iletim sırasında, klinik olarak kaydedilmiş EKG sinyalleri genellikle güç hattı girişimi, temas gürültüsü, kas kasılma gürültüsü, elektrot hareket artefaktı, taban çizgisi kayması ve sinyal işleme artefaktı gibi gürültü ve artefaktlarla bozulur. Mekansal, zamansal ve frekans alanında kaydedilen EKG sinyali ile seslerin etkileşimi istenen sinyali gizler ve ondan alınabilecek bilgi miktarını azaltır, klinisyeni yanlış yönlendirir ve tanı sürecini engeller. Gürültü kaynaklarını azaltmak, çıkarmak veya ayırmak için birçok filtreleme tekniği önerilmiştir, ancak sinyalin morfolojik şeklini bozmadan gürültüyü tamamen ortadan kaldıran sağlam ve güvenilir bir filtreleme tekniği bulmakta zorluk vardır. Son zamanlarda, Kör Kaynak Ayırma (BSS) ve Model Tabanlı Filtreleme teknikleri, biyomedikal sinyal işleme alanında iyi sonuçlar göstermiştir. Bu çalışmada, ölçülen EKG sinyal karışımını istatistiksel özelliklerine dayanarak bağımsız bileşenlere ayırmak için Bağımsız Bileşen Analizi (RobustICA) kullanılmış, daha sonra temiz bir sinyali tahmin etmek için bir tahmin algoritmasından yararlanılmıştır. Diğer taraftan, gürültüsüz bir sinyali tahmin etmek için Genişletilmiş Kalman Filtresi (EKF) ile doğrusal olmayan ECG sinyalinin doğrusal olmayan modelini kullanarak model tabanlı filtreleme gerçeklenmiştir. Önerilen algoritmalar, taban çizgisi kayması (BW), kas artefaktı (MA) ve elektrot hareketi (EM) gibi basit filtrelerle kaldırılamayan en yaygın EKG gürültülerini gidermek için kullanıldı. Önerilen yöntemlerin etkinliği, normal EKG sinyaline yapay olarak taban çizgisi kayması, kas artefaktı ve elektrot hareketi artefaktları eklenerek üretilen çeşitli gürültülü EKG sinyalleri 20 ila -20 Desibel arasında sinyal / gürültü oranı (SNR) olan bir veritabanı oluşturulmuştur. Filtreleme işleminden sonra filtrelenmiş EKG kayıtlarını SNR ve morfolojilerini incelenmiştir. ICA ve EKF'nin performans karşılaştırması, ICA'nın bazal dolaşımı ve elektrot hareketini azaltma oranını azaltma oranına kıyasla kas artefaktını azaltmada daha iyi sonuçlar ürettiğini gösterirken, EKF, sadece SNR düşük olduğunda kas yapısının azaltılmasında iyi sonuçlar gösterir. Analiz sonuçları, ICA'nın bazal dolaşan ve elektrot hareketi yapılarının azaltılmasında EKF'den daha iyi olduğunu gösterirken, her ikisi de kas artefaktının azaltılmasında iyi sonuç verdiğini göstermektedir.

CHAPTER 1. INTRODUCTION

The Electrocardiogram (ECG) is a record of the action potential differences originated from the heart's muscle contraction from the body surface. ECG signal shows the heart's rhythm, electrical conduction paths abnormalities and chambers state. The clinicians use the electrocardiogram as an essential part of the evaluation of the human heart functions because of the relative ease of acquisition, analysis and the valuable information it gives about the heart abnormalities. The normal ECG signal consists of a series of a repeated ECG cycle, each cycle described by waves, segments, and intervals. The clinicians need to identify the cardiac cycle components in order to give an accurate diagnosis.

The steps of ECG signal analysis are signal pre-processing and noise filtration using biomedical signal processing techniques, detection of the cardiac cycle component and feature extraction and formulation of the feature set [1]. The first step of ECG analysis is noise and artifacts removal because they can be within the frequency band of ECG and have the same shape. ECG contaminates with noise during the acquisition process as result of the muscle contraction, motion of the patient or electrodes and respiration. Also, the presence of electrical equipment around the patient represents another source of interferences. Filters detect, extract or separate the desired signal from a noisy background and reduce noise components.

Filtering techniques has been classified to linear and nonlinear filtering techniques. Wiener and Wavelet Filter are linear filtering techniques used when both the signal and noise within the same regions in the frequency domain, while model-based filtering and Blind Source Separation (BSS) are nonlinear techniques so do not depend on the linearity constraint [2].

Over the past several years, numerous research has been done in the area of ECG signal analysis and processing. However, recent advances in filtering, pattern recognition, and classification techniques have represented a new researching area in ECG signal analysis. Recently, Blind source separation techniques along with model-based filtering methods have shown good results in ECG signal processing.

1.1. Problem Statement

Finding robust and reliable ECG signal filtering techniques to remove muscle contraction artifact, baseline shift, and electrode motion artifact.

1.2. Objectives

In this study, we aim to use Independent Component Analysis and Extended Kalman Filter to filter a noisy ECG signal. Firstly, we developed a method based on RobustICA to separate observed source mixture into independent components and estimate a noise-free ECG signal. Secondly, we used Extended Kalman Filter with the nonlinear state-space model of ECG signal to estimate a noise-free signal. Finally, we evaluated the efficiency of the proposed methods in removing common ECG noises.

1.3. Thesis Layout

In this chapter besides the problem statement and the objectives, a general introduction has also been given; the following chapters of the thesis are organized as follow. Chapter 2 gives a background to the human heart anatomy, the human heart electrical system, the electrocardiogram wave description and acquisition, and the noise sources in the ECG signal. Chapter 3 reviews the literature of ECG signal filtering. Chapter 4 introduces the Blind Source Separation algorithms, which used mostly for ECG filtering. Chapter 5 presents the generation of ECG model and model-based filtering using Extended Kalman Filter. Following chapter describes the steps of the research methodology. Chapter 7 summarizes the results of the presented work. The conclusion of the whole work is in Chapter 8.

CHAPTER 2. HEART AND ELECTROCARDIOGRAM

Before starting any step of ECG signal analysis steps it is important to first understand the anatomy of the human heart and the physiology of ECG signal, to review the lead system which is used for signal acquisition, and to study the characteristics of noise and artifact sources.

2.1. The Human Heart Anatomy

The heart is a group of muscles, which contract rhythmically to push the blood all over the whole body. The blood provides the tissues of the body by oxygen and nutrition, and removes carbon dioxide and waste from the tissues to the kidneys and lungs for excretion. The heart is separated into four chambers. The right side contains two chambers are named right atrium and right ventricle, while the left side contains two chambers are known as left atrium and left ventricle.

The right chambers form a pump that circulates blood to the lungs. The right upper chamber accepts oxygen-poor blood coming from the body's tissues through the superior and inferior vena cava, and then it pushes the received blood through the tricuspid valve to the right bottom chamber (right ventricle) which contracts in order to supply the lungs with blood, that release carbon dioxide and relieves it with oxygen. Similarly, the left side forms a pump that pushes oxygen-rich blood to the body's organ. Firstly, the left upper chamber fills up with the blood coming from the lungs by the pulmonary vein, and then it pushes the blood through the mitral valve to the left lower chamber. The forceful contraction of left ventricle pumps the blood through the aorta to the body organs. Figure 2.1. shows the anatomy of the human heart, the heart's chambers, and valves.

The heart's muscle contract and relax rhythmically to supply the body with blood overall time. The contraction of the upper chambers supply the lower chambers with blood, while the forceful contraction of the lower right and left chambers send the blood to the lungs and to the body organs, respectively. The contraction period of the ventricles is called systole while its relaxation duration is known as diastole.

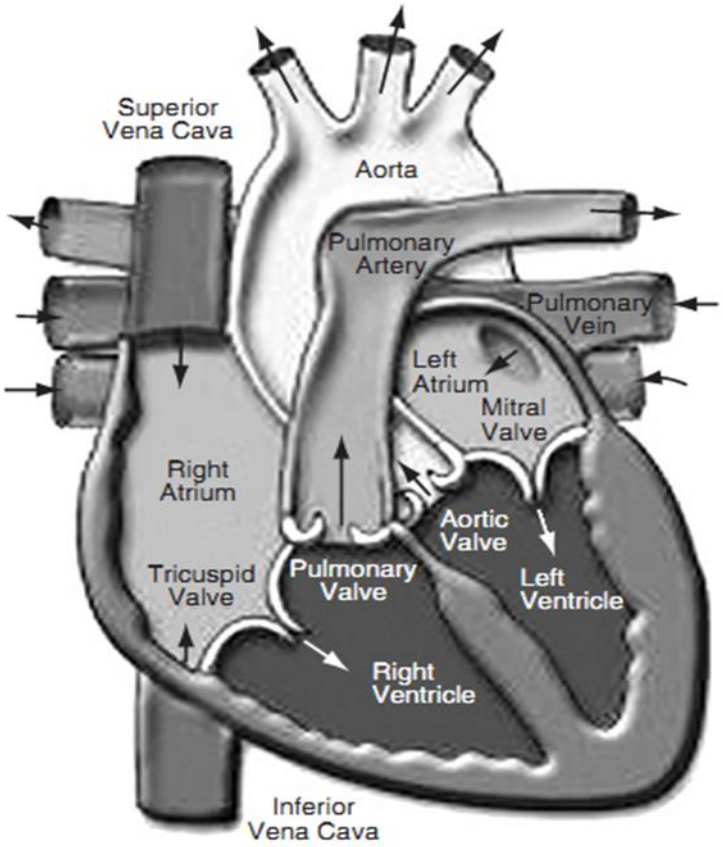

Figure 2.1. The anatomy of the human heart[1].

2.2. The Human Heart Electrical System

The electrical system of the heart is a group of specialized muscle located on the heart walls that maintain the cardiac cycle by generating electrical impulses to trigger the heart muscles to contract. The cardiac electrical system consists of the sinoatrial node (SA), the atrioventricular node (AV), bundle of His, and Purkinje fibers.

The sinoatrial node is a group of pacemaker cells placed in the upper wall of the right atrium, which generate a sequence of electrical impulse periodically without any external stimulus causing the atrial muscles to contract.

The atrioventricular node lies between the atrial septum and the ventricular septum in the right side. The AV node transfers the electrical impulse generated in SA node to the ventricles then, the impulse propagates to the Purkinje fibers through the bundle of His and, causing the ventricles to contract.

2.3. Electrocardiogram

The electrocardiogram (ECG) is a recorded graph to describe the action potential differences produced by depolarization and repolarization of the atria and the ventricle measured on the body surface using electrodes which, record the signal amplitude versus time. Figure 2.2. shows basic ECG waveform.

2.1.1. ECG wave description

ECG signal is a real-time signal, it is morphological features are described by waves, segments, and intervals as shown in figure 2.2. The letters; P, describes the contraction of the atrial, QRS represent the contraction of the ventricular, and T, describes the relaxation of ventricular, mark the locations of ECG waves.

Segments are known as the time durations between waves, normal ECG contains two segments PR the time duration between the P wave offset and R wave onset and ST the time duration between the S wave offset and T wave onset.

Intervals are known as the periods, which contain specific waves and segments. Normal ECG contains three intervals PR, ST, and QT. PR interval includes the time duration of the P wave and the PR segment. ST interval includes the time duration of the T wave and the ST segment. QT interval includes the time duration of the T wave, QRS complex, and the ST segment [1].

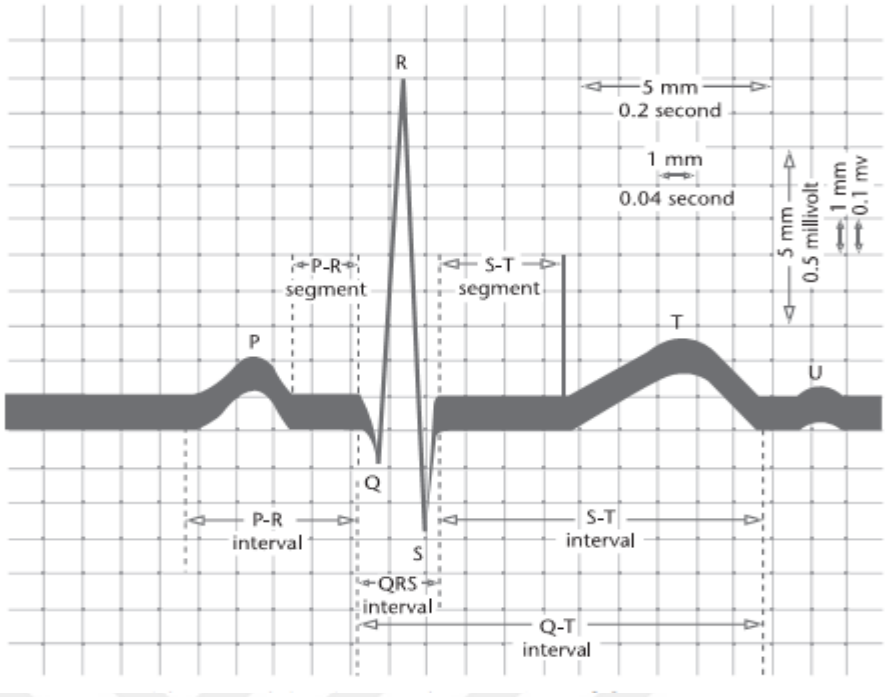

Figure 2.2. Normal electrocardiogram [2].

2.1.2. Leads system

The detection of the cardiac signal is achieved by using a non-invasive electrode placed on the skin of the patient at a specific point around the heart. Surface electrodes detect the action potential differences generated by the depolarization of excitable myocardium. The standard ECG records by using 10 electrodes which gives 12 different angles to view the heart. Each view is defined as a lead.

Electrodes according to the position of attachment on the body surface are labelled as follow:

The electrode, which is positioned on the left arm, is labelled as LA.
The electrode, which is located on the right arm, is labelled as RA.
The electrode, which placed on the left leg, is labelled as LL.
The electrode, which is positioned on the right leg labelled as RL (reference electrode).
The electrodes, which are placed on the chest, are labelled as V1 to V6.

In bipolar or Limb leads the electrodes are placed on the limbs. The views (I, II, III) are obtained using only two electrodes one performs as a positive electrode and the other is the negative electrode, Figure 2.3. shows electrode placements.

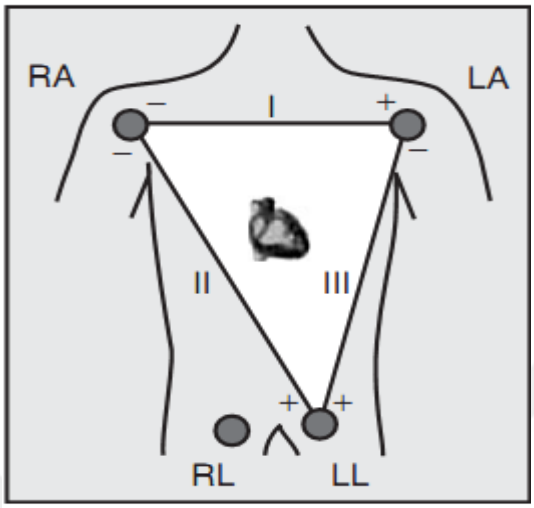

Figure 2.3. Electrode placements for the bipolar leads [1].

In unipolar leads system, the electrodes are located similarly to the unipolar leads to give three views (aVF, aVL, aVF). The views are obtained using three electrodes; one is positive while the average of the signals from the other two electrodes is negative. Figure 2.4. Shows electrode placements.

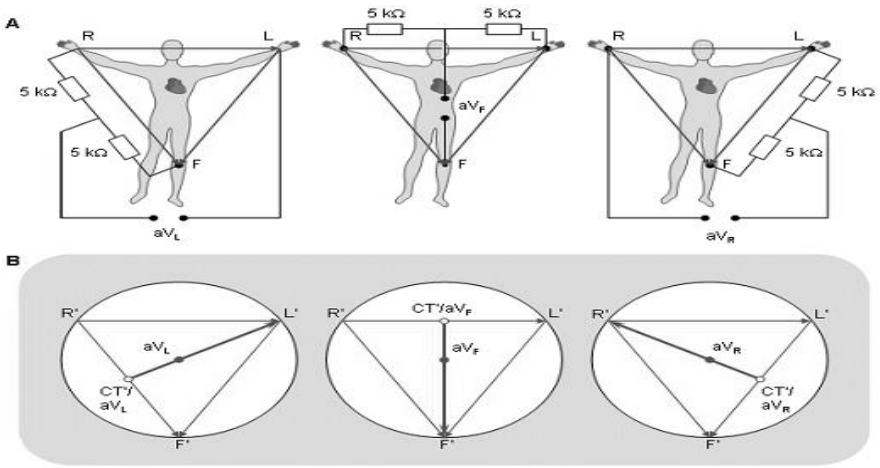

Figure 2.4. The unipolar leads calculation [1].

Unipolar precordinal leads (V1-V6) are six electrodes placed on the chest to give six views. The views are obtained using two electrodes one of the chest electrodes as positive while the negative electrode is the averaging of LA, RA and LL signals. For details see figure 4.6.

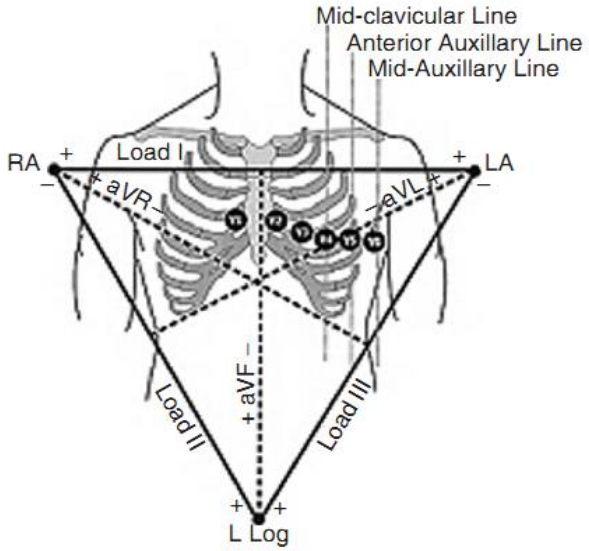

Figure 2.5. The unipolar chest electrodes placement [1].

2.4. Sources of Noise and Artifact in Records of Electrocardiogram

The recorded electrocardiogram contaminates with various kind of noises which have different frequency ranges during the acquisition process, these noises can interfere with ECG in the frequency domain, besides having the same morphological shape of ECG [3].

1. Power line interference(PLI) during biopotential measurements the cables which carry the recorded signals to the monitoring equipment are prone to electromagnetic interference with amplitude 50% of ECG amplitude and frequency (50 or 60 Hz) pickup and it is harmonics.

2. Baseline shift (BW) disturbance is produced from the variations in the electrode to skin impedance, or by patient's movements or by respiration movement with a frequency range between 0.1-0.3 Hz.

3. Electromyogram (EMG) noise is a surface biopotential signals generated from the muscles around the heart. These signals interfere with recorded ECG and could mask or alter its shape.

4. Contact noise is an unexpected change in ECG signal amplitude caused by the loss of contact between the body surface and electrode.

5. Patient electrode motion artifacts are changes in the resistance between surface electrodes and patient skin caused by patient movement, or respiration causing changes in ECG signal.

CHAPTER 3. LITERATURE REVIEW

Over the past years, several filtering methods have been proposed for detection, extraction or separation of the desired ECG signal from a noisy background and reduce noise components. In this chapter, we will discuss the old and recent research in ECG filtering. We will focus on the role of blind source separation techniques and model-based filtering in ECG noise removal as the most recent filtering techniques.

Many of the avaliable filtering techniques are based on the notation of spectral decomposition of the signal [4]. Such techniques include notch filters for removing the effect of the electrical mains supply, and both low and high bandpass filters for removing noise that is localized in particular regions of the frequency spectrum. These techniques all depend on the principle of linear superposition and there is a fundamental assumption that the underlying signal and the noise are active in different parts of the frequency spectrum. Finite Impulse Response filters are used in all ECG devices to filter undesired low and high frequency noises, and removal or reduction of baseline wanders and power line interference. The ECG signal frequency band is between 0.01-100 Hz, signals with frequencies greater than 100 Hz or less than 0.01 Hz are noise signals. Low and high pass filter is implemented to filter the high and low frequency noise signals. The frequency of power line is 50 or 60 Hz, which is within the frequency range of ECG signal stop filter so a selective filter is used to remove it [3].

Filtering techniques are classified to linear and nonlinear filtering techniques. Wiener Filtering and Wavelet Filtering are Linear filtering methods used when both the signal and noise are within the same regions in the frequency domain. Nonlinear noise reduction (NNR), independent component analysis (ICA) ,and model-based filtering are nonlinear techniques do not depend on the linearity constraint [2].

Wiener filter is the first Model-Based filter in the frequency domain that predicts a noise-free signal based on an estimated model of signal and noise. Wiener filter assumes that the signal and noise together are not deterministic signals; this diminishes the performance of the optimal Wiener filter in ECG filtering since ECG signal has some deterministic element. [2].

Within the last 30 years, several filtering techniques have been developed based on Wiener and Wavelet Filtering Method. An adaptive Wavelet Wiener Filter has been proposed to remove myopotentials (EMG) noise and, to estimate a noise-free ECG signal in [5]. The method has shown good result when tested on a generated database using signals from the CSE database and compared with other Wiener and Wavelet based Filtering Method. An Improved Wavelet Wiener Filter has also been proposed to overcome the linear filtering problem [6].

Blind source separation techniques attempting to separate the signal from noise mixture relying on the statistical features of the data. On the other hand, Model-based filtering attempt to build an approximate model of the signal or noise into the filtering system, so a more successful filter can be developed. Developing a dynamics model of ECG signal and fitting it to a section of ECG can create a de-noised ECG signal.

Over the past several years, source separation methods have gotten a lot of consideration for their ability to separate noise sources from non-Gaussian noisy ECG signal and make it cleaner and better interpretable for the clinician. The ICA technique works on maximizing non-Gaussianity using higher order statistics such as kurtosis and Negentropy. The maximization process is based on the central limit theorem [7].

ECG signal has non-Gaussian shape, therefore [8] has shown the ability of FASTICA to remove baseline wandering from a single channel ECG which has been constructed to a multi-channel by adding some delay to the original signal. Results derived by the proposed method were compared with those obtained from traditional FIR high-pass filtering. Different ICA algorithm such as JADE and FASTICA have been studied and used for ECG de-noising, with an effort to use constraintICA for noise separation in.

Signals from MIT-BIH database have been mixed randomly then ICA algorithm has been executed to separate signals [9].

FASTICA with two granularity function; hyperbolic tangent function and power have been applied to separate noise components from noise and signal mixture. Fast ICA has been used on data taken from MIT-BIH arrhythmia database, Analysis of the results has shown that tanh granularity function has quicker convergence when matched with multi-granularity power function as in [10].

There are number of approaches to achieving artifact removal using kalman filter, firistly, A dynamical model for producing artificial electrocardiogram signals has been introduced in [11] based on three joined conventional differential equations with the ability of the operator to specify the model parameters. A technique for simultaneously filtering, compressing and classifying the ECG has been described. The filtering process has been accomplished by fitting a set of six Gaussians, each specified by three parameters in an ordinary differential equation, and performing a constrained nonlinear optimization, this paper demonstrate that in-band noise can be completely removed [12]. Another one propose a mathematical framework for the model-based Bayesian filtering of single channel noisy ECG recordings. Within this framework several suboptimal filtering schemes were developed and the results were compared with conventional filtering methods [13].

A new EKF algorithm have been introduced to suppress the noise of electrocardiogram signal using the parameters of ECG dynamical model. The EKF structure has been constructed by introducing a simple autoregressive (AR) model for each of the 15 dynamic parameters of the Gaussians [14], [15]. These works has controbuted in the introduction of a wave-based state space formulation for generating different types of pathologic ECG waveforms. The derivation of a linearized model and the establishment of linear observation relations, using a Kalman-based filtering scheme that could provide robust estimations of the input noisy measurements and the determination of ECG fiducial points based on the wave-based structure [16].

A modified EKF structure has been designed in [17] for filtering the cardiac signals. The designed EKF dynamic model depended on the three-dimensional nonlinear model. This nonlinear model has been linearized in order to apply the proposed EKF. The structured filter has been used to filter ECG signals, and the obtained results illustrate the filter is able to follow and remove the noise components.

Accurate metrics for evaluating the effectiveness of filtering techniques applied to the ECG are difficult to define due to the inherently complicated structure of the noise and the absence of knowledge about the underlying dynamical processes. Without having access to a noise-free ECG signal, the fact that the true underlying dynamics of a real ECG can never be known implies that one cannot distinguish between the clean ECG signal and the many sources of noise that can occur during recording in a clinical environment. While the availability of biomedical databases provides a useful benchmark for comparing different techniques, this approach can never truly distinguish between noise and signal. ECGSYN, a dynamical model for generating ECG signals [18] with known temporal and spectral characteristics and prespecified average morphology is used to compare and evaluate these techniques under a range of different conditions.

CHAPTER 4. THEORY

In the first part of this chapter, the basics of independent component analysis and some of its algorithim will be introduced. While in the second part, the extended Kalman filter and the dynamical ECG model will be explained.

4.1. Independent Component Analysis

ICA is a nonlinear statistical analysis for separating a set of signals depending on their structure into statistically independent components.

4.1.1. The basic concept of ICA

Independent Component Analysis is a source separation technique that transforms the data onto an independent set of vectors. ICA is a biorthogonal transformation project the data mixture into non-orthogonal axes, as opposite to Principal Component Analysis which is an orthogonal transformation [7]. The following general mathematical framework is used to define ICA.

$$X = AS \tag{4.1}$$

Where X is the observed signals matrix, A is unknown mixing matrix, and S is the unknown source signals matrix.

$$S = A^{-1} X = W X \tag{4.2}$$

W Is a demixing matrix that is obtained by inversing the matrix A.

ICA assumes that the observed signals are linear mixtures of independent source signals and source signals are non-Gaussian to estimate the unknown source signals. The general principles of source estimations are nonlinear decorrelation, if the sources are uncorrelated then they are independent, or by maximization of source non-gaussianity. The signals non-Gaussianity can be measured by maximum likelihood, mutual information, marginal entropy, negentropy, and kurtosis [7].

The kurtosis is used usually to measure non-gaussianity for its computational and theoretical simplicity, which is the normalized fourth-order cumulant when the data is pre-processed so that its variance is equal one $E\{S^2\} = 1$.

$$kurt(S) = E\{S^4\} - 3(E\{S^2\})^2 \tag{4.3}$$

The data need to be centered and whitened before applying ICA as pre-processing step to simplify ICA algorithms.

- Centering: is a simple preprocessing step performed by subtracting the observed data from it is mean $\bar{X}$, centering can be expressed by:

$$X_c = X - \bar{X} \tag{4.4}$$

- Whitening: is performed to remove the linear dependencies between the observed data, which mean the data are uncorrelated and their variances equal unity, whitening can be expressed by:

$$Z = VX = VAS \tag{4.5}$$

Z is the whitened data and V is the whitening matrix.

4.1.2. ICA algorithms

Various ICA algorithms have been developed during the last years but in this section we will focus on the algorithms which are used mostly in noise removal from ECG, which are summarized as follow:

4.1.2.1. JADE algorithm

JADE algorithm has been developed and implemented by (Cardos, 1993), is built depending on joint diagonalization of matrices. JADE algorithm uses second order moment to perform the whitening step, which produces the whitening matrix and whitened sources while it uses fourth order moment in order to perform the separation of the source signals from the observation mixture.

4.1.2.2. FastICA algorithm

FastICA is a fixed-point algorithm has been developed by (Hyvärinen, 1999), it is another source separation algorithm works on non-Gaussianity maximization as a measure of independence between the source components. FastICA algorithm has been built depending on the central limit theorem, which states: "The distribution of a sum of independent random variables tends toward a Gaussian distribution"[19].

Steps of applying FastICA algorithm is summarized as follow.

- Data pre-processing: Center and Whiten the data to make it zero mean and unit variance.
- Set i=0 and give the de-mixing vector w an initial value.
- Correct w :$w(i+1) = E\{zg(w_i^T z)\} - E\{zg'(w_i^T z)\}w_i$
- Normalize w: $\mathrm{w(i+1)} = \frac{w(i+1)}{\|w(i+1)\|}$
- Check the convergence and continue until converge.
- Calculate the estimated component value using $S_1 = wZ$.

$Z = z_1, z_2, \ldots z_n$ is whitened data matrix and $S = y_1, y_2, \ldots y_n$ are the estimated data matrix.

4.1.2.3. RobustICA algorithm

RobustICA is one of independent component analysis developed by (V. Zarzoso and P. Comon , 2010) as an improvement of FastICA. The algorithm achieves a precise line search of the absolute fourth order moment to extract any nonzero independent component. In this method, the kurtosis divergence function is optimized depending on an optimal step size searching method. This technique calculates the step size globally at each de-mixing vector update to enhance the kurtosis in the search path [20],[21].

RobustICA implements optimization at each iteration using an optimal step-size method; Steps can be summarized as follow

- Set i=0 and give the de-mixing vector w an initial value.
- Calculate the optimal step-size polynomial coefficients.$p(u) = \sum_{k=0}^{4} a_k u^k$ The coefficients a_k can simply be acquired from the observed signal and the recent values of w and g.
- Compute the fourth order polynomial roots $\{u_K\}_{k=1}^{4}$.
- Choose the root which maximize the absolute value of fourth order moment within search path using:
  $u_{opt} = \max_k |kurt(w + u_k g)|$
- Update $w(i + 1) = w + u_{opt} g$.
- Normalize $w(i + 1) = \frac{w(i+1)}{\|w(i+1)\|}$.
- Calculate the estimated component value using $S_1 = wZ$.

While g is the search direction and typically the gradient $g = \nabla_w kurt(w)$.
$S = y_1, y_2, \ldots y_n$ are the estimated data matrix.

4.2. Extended Kalman Filter

Kalman filter is a statistical method that optimally estimates past, present and future states of an unknown variable within a set of noisy variables. The extended Kalman filter (EKF) is a modified version from the traditional Kalman that has been developed to work on the dynamical nonlinear models [2]. For a discrete nonlinear model and observation vector y_k, can be formulated as follows:

$$\begin{aligned} x_{k+1} &= f(x_k, w_k, k) \\ y_k &= g(x_k, v_k, k) \end{aligned} \tag{5.1}$$

Where f is the state space update function with state vector x_k , w_k represents the process noise vector with associated covariance matrix $Q_k = E\{w_k w_k^T\}$. While g introduces the measurement function, which shows the relationship between the state vector and the observations and v_k is measurement noise vector with associated covariance matrix $R_k = E\{v_k v_k^T\}$.

The initial state estimate vector p is also expected to be known and is calculated from$\bar{x}_0 = E\{x_0\}$, with $P_0 = \{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\}$.

The above equations (Equations 5.1) need to be linearized around a specific point$(\hat{x}_k, \hat{w}_k, \hat{v}_k)$ in order to compute the Kalman filter gain and the covariance matrix [2]. This leads to the following linear approximate model:

$$\begin{cases} x_{k+1} \approx f(\hat{x}_k, \hat{w}_k, k) + A_k(x_k - \hat{x}_k) + F_k(w_k - \hat{w}_k) \\ y_k \approx g(\hat{x}_k, \hat{v}_k, k) + C_k(x_k - \hat{x}_k) + G_k(v_k - \hat{v}_k) \end{cases} \tag{5.2}$$

$$\begin{cases} A_k = \left.\frac{\partial f(x, \hat{w}_k, k)}{\partial x}\right|_{x = \hat{x}_k} \quad F_k = \left.\frac{\partial f(\hat{x}_k, w, k)}{\partial w}\right|_{w = \hat{w}_k} \\ C_k = \left.\frac{\partial g(x, \hat{v}_k, k)}{\partial x}\right|_{x = \hat{x}_k} \quad G_k = \left.\frac{\partial g(\hat{x}_k, v, k)}{\partial v}\right|_{v = \hat{v}_k} \end{cases} \tag{5.3}$$

In order to have a simpiler matrix form, the matrices F_k and G_k can be mixed with the noise covariance matrices $Q_k \leftarrow F_k Q_k F_k^T$, $R_k \leftarrow G_k R_k G_k^T$. With these notations, the the time propagation and the measurement propagation equations may be introduced using the following equations:

$$\begin{cases} \hat{x}_{k+1}^- = f(\hat{x}_k^-, w, k)|_{w=\bar{w}_k} \\ P_{k+1}^- = A_k P_k^+ A_k^T + Q_k \end{cases} \quad (5.4)$$

$$\begin{cases} \hat{x}_k^+ = \hat{x}_k^- + K_k r_k \\ P_k^+ = P_k^- - K_k C_k P_k^- \end{cases} \quad (5.5)$$

$$K_k = P_k^- C_k^T \left[C_k P_k^- C_k^T + G_k\right]^{-1} \quad (5.6)$$

where r_k is the innovation signal

$$r_k = y_k - g(\hat{x}_k^-, v, k)|_{v=\bar{v}_k} \quad (5.7)$$

Where $\hat{x}_k^- = \hat{\mathbb{E}}\{x_k | y_{k-1}, \dots, y_1\}$ is an estimation of the previous states using the observations y_1 to y_{k-1}, $\hat{x}_k^+ = \hat{\mathbb{E}}\{x_k | y_k, \dots, y_1\}$ is an estimation of the coming states using the k-th observation ,and P_k^- is an estimation of the a previous covariance matrices using the observations y_1 to y_{k-1}. While, P_k^+ is an estimation of the coming states covariance matrices after using the k-th observation [15].

4.2.1. An artifical ECG generator

ECG is a series of a repeated super quasi-periodic waves (P, Q, R, S, T), which lies in (x, y)-plane. The ECG model generator uses a series of exponentials to trace out the morphology of the ECG in the z-direction.The periodicity of the ECG makes the progress of the ECG track in a bounded cycle of unit radius. The extrema of the peaks (P, Q, R, S, T) is described by (θ_P, θ_Q, θ_R, θ_S, θ_T) has proposed a synthetic ECG generator. The following differential equations has been developed in [18] are ECG dynamic model equations which are a set of three ordinary state in Cartesian coordinates .

$$\begin{cases} \dot{x} = \alpha x - \omega y \\ \dot{y} = \alpha x + \omega y \\ \dot{z} = -\sum_{i \in \{P,Q,R,S,T\}} a_i \, \Delta\theta_i exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) - (z - z_0) \end{cases} \quad (5.8)$$

$$\alpha = 1 - \sqrt{x^2 + y^2} \quad (5.9)$$

$$\theta = atan2(y, x) \quad (5.10)$$

$$\Delta\theta_i = (\theta - \theta_i) mod(2\pi) \quad (5.11)$$

Where x, y, and z are the state variables, ω is the angular velocity of the track, θ is the four quadrant arctangent of the components of x and y, with $-\pi \leq atan2(y, x) \leq \pi$ and z_0 is the baseline. θ_i is the angular position for the PQRST waves, a_i is the magnitude of the peaks, b_i is the width (time duration) of each peak. Plotting z coordinate vs. time presents the artificial ECG signal [22].

4.2.2. The polar form of ECG dynamic model

The Cartesian form equations (Equations 5.1) can be rewriten in the polar form as follow.

$$\begin{cases} \dot{r} = r(1 - r) \\ \dot{\theta} = \omega \\ \dot{z} = -\sum_{i \in \{P,Q,R,S,T\}} a_i \, \Delta\theta_i exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) - (z - z_0) \end{cases} \quad (5.12)$$

The polar form is simpler compared to the cartesian form, We can notice that r can be discarded since it doesn't affect the state variable z or any other variable.

The two-dimensional equations of the dynamic system can be simplified and written in the discrete form using a time step of size δt as follow.

$$\begin{cases} \theta_{k+1} = \theta(k) + \omega\delta t \\ z_{k+1} = -\sum_i \delta t a_i \Delta\theta_i exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) + z(k) + \eta\delta t \end{cases} \quad (5.13)$$

Where $\theta(k)$ and $z(k)$ are the phase and ECG at time instant k, η is an additive noise that describes all the additive sources of process noise and replaces baseline wander, i is the number of Gaussian functions while θ_i represents the phase center of the i-th Gaussian.

4.2.3. Linearization of ECG model

In nonlinear dynamic models, the EKF is used to change the states over time but to complete the estimation process the model needs to be linearized to compute the Kalman filter coefficients. For this, θ and z are considered as the state variables and the model parameters a_i, b_i, θ_i, ω, η are the process noises, By defining. In order to apply the EKF

$$\begin{cases} x_{k+1} = F(\theta, \omega, k) \\ y_k = G(\theta, z, \omega, a_i, \theta_i, b_i, \eta, k) \end{cases} \quad (5.14)$$

The linearized model with respect to θ and z is represented with the following equations:

$$\begin{cases} \frac{\partial F}{\partial Z} = 0 \qquad \frac{\partial F}{\partial \theta} = \frac{\partial G}{\partial Z} = 1 \\ \frac{\partial G}{\partial \theta} = -\sum_{i\in\{P,Q,R,S,T\}} \delta . a_i \left[1 - \frac{\Delta\theta_i^2}{b_i^2}\right] exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) \end{cases} \quad (5.15)$$

Similarly, the linearization of (Equations 5.14) with respect to the process noise components.

$$\begin{cases} \frac{\partial F}{\partial a_i} = \frac{\partial F}{\partial b_i} = \frac{\partial F}{\partial \theta_i} = \frac{\partial F}{\partial \eta} = 0 \quad i \in \{P, Q, R, S, T\} \\ \frac{\partial F}{\partial \omega} = \delta \end{cases} \quad (5.16)$$

$$\begin{cases} \frac{\partial G}{\partial a_i} = -\delta . \Delta\theta_i \, exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) \\ \frac{\partial G}{\partial b_i} = -\delta . a_i \frac{\Delta\theta_i^3}{2b_i^3} exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) \\ \frac{\partial G}{\partial \theta_i} = \delta . a_i \, [1 - \frac{\Delta\theta_i^2}{b_i^2}] exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) \\ \frac{\partial G}{\partial \omega} = \ 0 \end{cases} \quad (5.17)$$

The system process noise and covariance vectors are defined as follows:

$$\begin{cases} w_k = [a_P, \dots, a_T \, , b_P, \dots, b_T, \theta_P, \dots, \theta_T, \omega, \eta]^T \\ Q_k = E\{w_k w_k^T\} \end{cases} \quad (5.18)$$

The phase observations $\emptyset_k$ and the noisy ECG measurements s_k can be written in state space form as coming:

$$\begin{bmatrix} \emptyset_k \\ s_k \end{bmatrix} = \ \mathrm{I}. \begin{bmatrix} \theta_k \\ z_k \end{bmatrix} + \begin{bmatrix} u_k \\ v_k \end{bmatrix} \quad (5.19)$$

Where $R_k = E\{[u_k, v_k]^T [u_k, v_k]\}$ is the observation noise covariance matrix [2].

4.2.4. The extended kalman filter smoother

The extended Kalman smoother (EKS) uses the information of future observations to give better estimates of the current state. Due to this non-causal nature, the EKS is expected to have a better performance compared with the EKF. The EKS algorithm basically consists of a forward EKF stage followed by a backward smoothing stage.

CHAPTER 5. RESEARCH METHODOLOGY

The first step of ECG analysis is noise and artifacts removal because ECG contaminates with noise during the acquisition process as a result of the muscle contraction, motion of the electrodes, or patient respiration. These noises can be within the frequency band of ECG and have the same morphology.

Until now, numerous methods have been introduced for noise reduction. In spite of the numerous research in the area of ECG signal processing, finding robust signal filtering techniques still form a challenge because the success of the noise reduction algorithm does not rely only on the nature of the signal but also depends on the noise nature.

In this chapter, we propose to use independent component analysis and extended Kalman filter for muscle artifact, baseline shift, and electrode motion noises removal and estimation of a noise-free ECG signal.

5.1. ECG Filtering Using Independent Component Analysis

In this part, depending on the held comparison in [20] between newly proposed RobustICA and FastICA . RobustICA, which is one of Independent Component Analysis algorithms, has been used for muscle artifact, baseline shift, and electrode motion noises separation from contaminated ECG signal using only two leads. Estimation method based on correlation has been performed to reconstruct a clean signal from the separated ICs. Signals are taken from MIT-BIH database and mixed at a different ratio with noises taken from MIT noise stress test database.

$$\begin{cases} modified\ lead\ II = lead\ II + N \\ modified\ v5 = v5 + N \end{cases} \tag{6.1}$$

Where the modified leads represent the signal and noise mixture. $lead\ II$ And $v5$ are ECG signals taken from MIT-BIH database. N represents noise source (BW, MA, or EM) taken from MIT noise stress test database.

RobustICA algorithm has been used to separate the observed ECG signal and noise mixture into independent variable. Using visual inspection to differentiate the desired signal from noise component is not accurate. The statistical properties of separated independent components can be used in the recognition process. The kurtosis of normal ECG is bigger than the noise kurtosis so the component with the maximum kurtosis is selected to be ECG signal.

The steps of the algorithm can be described as follow.

- Input: Observed ECG signal
- Preprocessing: Subtract the signal mean.
- Prewhiten the signal.
- Estimate independent components using RobustICA.
- Estimate mixing matrix.
- Calculate Kurtosis of estimated independent components.
- Output: Estimated Filtered ECG signal.

5.2. ECG filtering using Extended Kalman Filter

In this part, we propose to use and evaluate the efficiency of Extended Kalman Filter for muscle contraction artifact, baseline shift, and electrode motion noises filtering from contaminated ECG signal. Using EKF formulation and the nonlinear state-space model presented in the previous chapter an estimation of the artifacts can be found.

Process Equation:

$$\begin{cases} \theta_{k+1} = (\theta_k + \omega\delta)\ mod(2\pi) \\ s_{k+1} = -\sum_{i}^{N} \delta ta_i\ \Delta\theta_i exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) + s_k + \eta \end{cases} \tag{6.2}$$

δ Is the sampling time,$\Delta\theta_i = (\theta_k - \theta_i) mod(2\pi)$, $\omega = 2\pi f$, f is heart rate, and N is the number of Gaussian functions.

$$\begin{cases} \emptyset_k = \theta_k + u_k \\ s_k = \ z_k + v_k \end{cases} \quad (6.3)$$

Where $\emptyset_k$ is the phase of the observations, and s_k is the measured noisy ECG. While u_k and v_k are the measurement noises

The noise estimation process using EKF can be summarized in the following scheme.

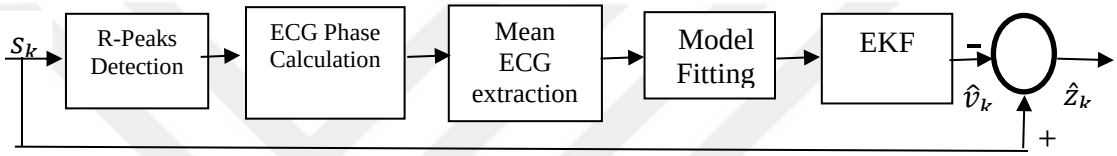

Figure 5.1. The overall de-noising scheme using EKF.

The estimated de-noised ECG signal can be found from the following equation:

$$\hat{z}_k = s_k - \hat{v}_k \quad (6.4)$$

Where $\hat{z}_k$ is the estimated de-noised ECG signal, s_k is the noisy ECG signal and $\hat{v}_k$ is the estimated noise.

- R-peaks detection and phase calculation

Peaks are required for constructing the phase signal $\emptyset_k$, which is in terms needed for synchronizing the noisy ECG with the dynamic model. They are also used for extracting the mean contaminating signals by synchronous averaging over the heart beats. The R-peaks may be detectable from the noisy recordings or from any arbitrary ECG channel synchronously recorded with the noisy dataset by using peaks detection algorithm.

- Mean ECG extraction

Using the R-peaks, the ensemble average (EA) and standard deviation of the contaminating signals are extracted through synchronous averaging. In order to automate the parameter selection procedure for any given ECG, the parameters should be estimated from the signal itself. For this, any noisy ECG transformed to a three-dimensional representation by plotting the noisy ECG versus the periodic phases that are assigned to each sample in polar coordinates on the unit circle (r =1).

- Model fitting

In this step, the problem is to find the optimal parameters that can best fit the mean ECG, by using a nonlinear least square estimation, the best parameters of the Gaussian kernel in the MMSE senseare found, such that the model will best fit the mean contaminating signals waveform.
A practical means of solving this nonlinear leastsquares problem is the lsqnonlin function of Matlab® that was used to fine tune the parameters which gives also minimum model error.

- Filtering

Having extracted all the required parameters, the contaminating signals may be estimated by the KF framework and the desired background signal is found.

CHAPTER 6. RESULTS

With the aim of assessing the efficiency of the suggested methods in muscle contraction artifact (MA), baseline wandering (BW), and electrode motion (EM) noises reduction an artificial database has been generated by selecting some files from MIT-BIH Arrhythmia database and mixing the normal ECG signals with a noise signals from MIT-BIH non-stress test database. The selected files have been segmented so that every 10 seconds represent a signal, then 7 combinations with a different signal to noise ratio values extending from -20 to 20 dB is acquired for each signal. The signals mixing process has been accomplished by using an algorithm to generate a desired noisy signal and adding it to clean ECG. The signal to noise ratio of the output signals is then calculated and compared to the reference noise-free signal. In this chapter, the x-axis in the entire figure is time in second (s) and the y-axis is amplitude in millivolt (mv). All the signal is 10s length or 3600 sample. The entire testing has been simulated using Matlab® source codes curried out on a i3-2350M CPU 2.30 GHz.

6.1. Databases

- MIT-BIH Arrhythmia database, which contains 48, file two channel ECG.
- MIT-BIH non-stress test database (NSTDB), which contains a real ECG noises.

6.2. Noise Generation

Muscle contraction artifact, baseline wandering, and electrode motion noises are the most common noise sources in ECG signal that the simple filtering method fails to efficiently remove them.

6.2.1. Baseline wandering (BW)

Baseline wandering noise usually originates from the patient's movements or respiration, the following figure shows the shape of the baseline disturbance. The signal has been taken from NST Database and plotted using MATLAB.

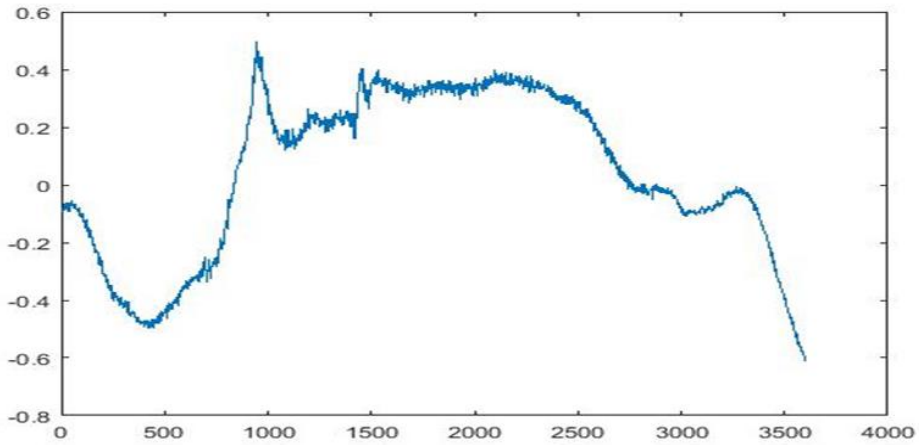

Figure 6.1. Baseline wander noise.

6.2.2. Muscle artifact (MA)

Muscle artifact is an bioelectrical distubance orginate from the contraction of muscle the following figure shows the shape of the muscle contraction disturbance. The signal has been taken from NST Database and plotted using MATLAB.

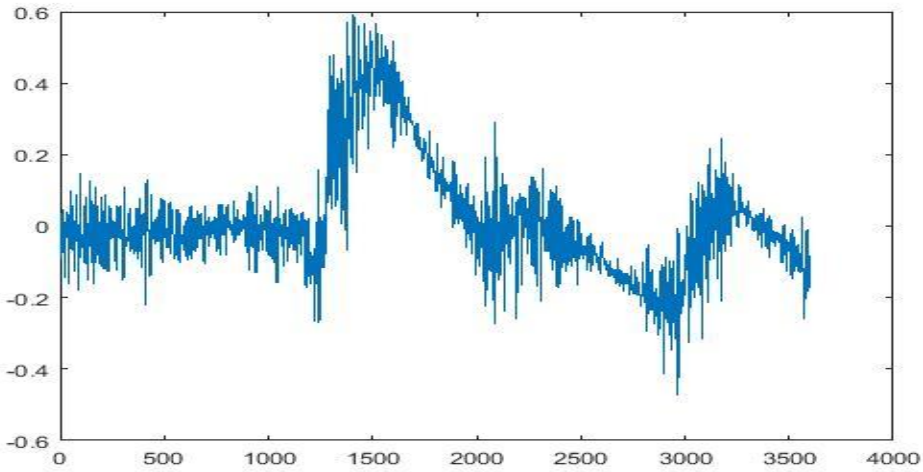

Figure 6.2. Muscle artifact.

6.2.3. Electrode movement (EM)

Electrode movement artifact generates from the contact loosing between the electrode and patient skin, the following figure shows the shape of the electrode movement disturbance. The signal has been taken from NST Database and plotted using MATLAB.

.

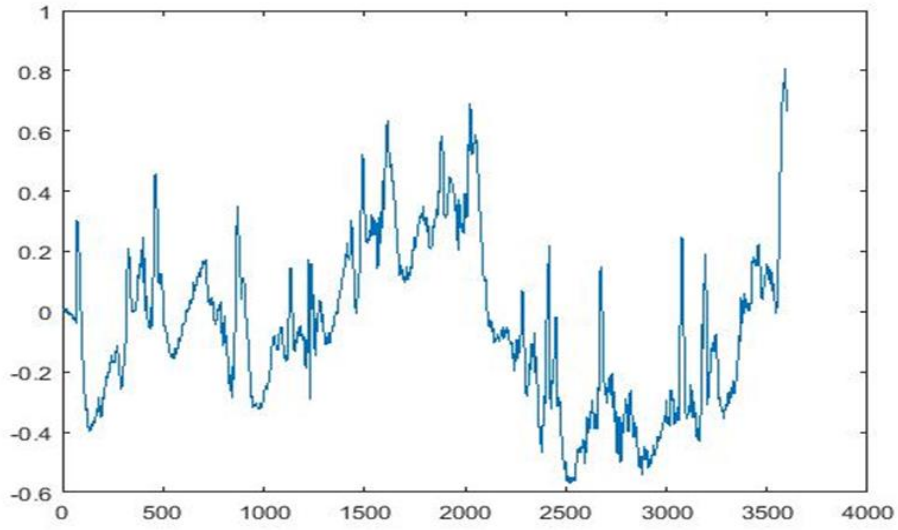

Figure 6.3. Electrode movement artifact.

6.3. Evaluation Metrics

6.3.1. Root mean square error (RMSE)

RMSE is an evaluation method used to assess the proposed method by calculating the square root of error mean RMSE is calculated from:

$$\text{RMSE} = \sqrt{\frac{1}{K}\sum_{k=1}^{K-1}(S_k - SE_k)^2} \qquad (7.1)$$

Where S_k is original signal SE is the filtered estimated signal and K is the samples number. The small RMSE value indicates that the filtering technique minimizes the noise with less distortion.

6.3.2. Signal to noise ratio (SNR)

$$SNR(DB) = 20log\frac{rms(signal)}{rms(noise)} \quad (7.2)$$

SNR is opposite to RMSE the high values of SNR means that filtering method minimizes the noise with less distortion [2].

6.4. Results of applying RobustICA:

In order to evaluate the efficiency of RobustICA in ECG noises reduction a file 100 from MIT-BIH Arrhythmia database has been selected and combined with baseline wandering , EMG artifact , and electrode motion noises taken from MIT-BIH non-stress test database. The selected files has been segmented and mixed to produce a noisy signal have signal to noise ratio values between 20 to -20 dB. The following tables and figures illustrate the signal to noise ratio before and after using RobustICA algorithm.

6.4.1. Baseline wander filtering using RobustICA

Baseline wandering noise has been mixed with a normal ECG signal to obtain noisy signals with signal to noise ratio (20, 12, 6, 0, -6, -12, -20) dB. The following figures show the noisy signal and the filtered signal using RobustICA algorithm.

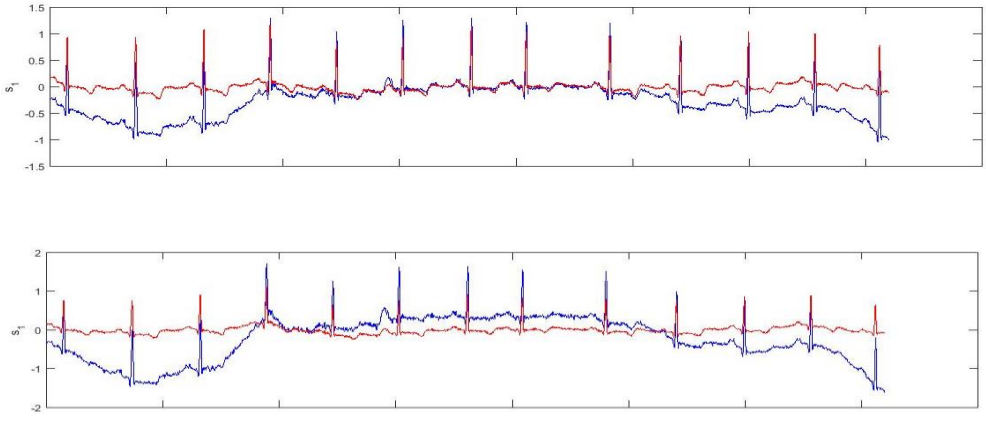

Figure 6.4. Baseline wander filtering using RobustICA

In Figure 7.4. the top panel, the blue line shows ECG signal combined with baseline wandering noise (SNR =-6 DB) and the red line shows the estimated noise free ECG signal. In the bottom panel the blue line shows ECG signal combined with baseline wandering noise (SNR= 0 DB) and the red line shows the estimated noise free ECG signal.

Table 6.1. Results of baseline wander removing using RobustICA.

Baseline wander (BW)					
SNR DB before	SNR DB after	RMSE	RMSEDB	kurtosis of c1	kurtosis of Ic2
-20	-7.8257	9.2261%	-10.3498	4.7999	-1.2888
-12	0.17434	8.8305%	-10.5402	8.7293	-1.2284
-6	6.1743	8.476%	-10.7181	16.1427	-1.1815
0	12.1743	8.2784%	-10.8205	21.4774	-1.0092
6	18.1743	8.1726%	-10.8764	23.5708	0.21757
12	24.1743	7.9883%	-10.9755	24.4973	5.4882
20	32.1743	7.6255%	-11.1773	31.9039	16.3704

6.4.2. Electrode motion artifact filtering using RobustICA

Electrode motion noise has been mixed with a normal ECG signal to obtain noisy signals with signal to noise ratio (20, 12, 6, 0, -6, -12, -20) dB. Figure 7.2. shows the noisy signal and the filtered signal using RobustICA algorithm.

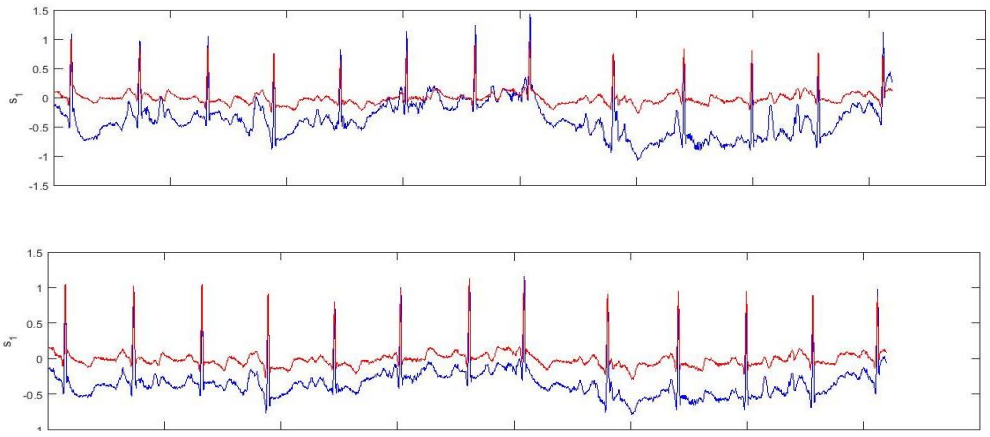

Figure 6.5. Electrode motion artifact filtering using RobustICA

In Figure 7.5. the top panel the blue line shows ECG signal combined with electrode motion artifact (SNR =0 DB) and the red line shows the estimated noise free ECG signal. In the bottom panel, the blue line shows ECG signal combined with electrode motion artifact (SNR= 6 DB) and the red line shows the estimated noise free ECG signal.

Table 6.2. Results of electrode motion removing using RobustICA.

Electrode movement (EM)					
SNR DB before	SNR DB after	RMSE	RMSE DB	Kurtosis of Ic1	Kurtosis of Ic2
-20	-6.5408	9.324%	-10.304	-1.005	0.55617
-12	-0.7976	9.1358%	-10.3925	1.4592	2.171
-6	7.4592	8.8052%	-10.5526	7.5852	-0.76359
0	13.4592	8.473%	-10.7196	16.5418	-0.35156
6	19.4592	8.258%	-10.8312	21.7719	2.3317
12	25.4592	8.0865%	-10.9224	23.6207	9.4536
20	33.4592	7.5933%	-11.1957	30.7979	16.7482

6.4.3. Muscle artifact filtering using RobustICA

Muscle contraction artifact has been mixed with a normal ECG signal to obtain noisy signals with signal to noise ratio (20, 12, 6, 0, -6, -12, -20) dB. Figure 7.3. shows the noisy signal and the filtered signal using RobustICA algorithm.

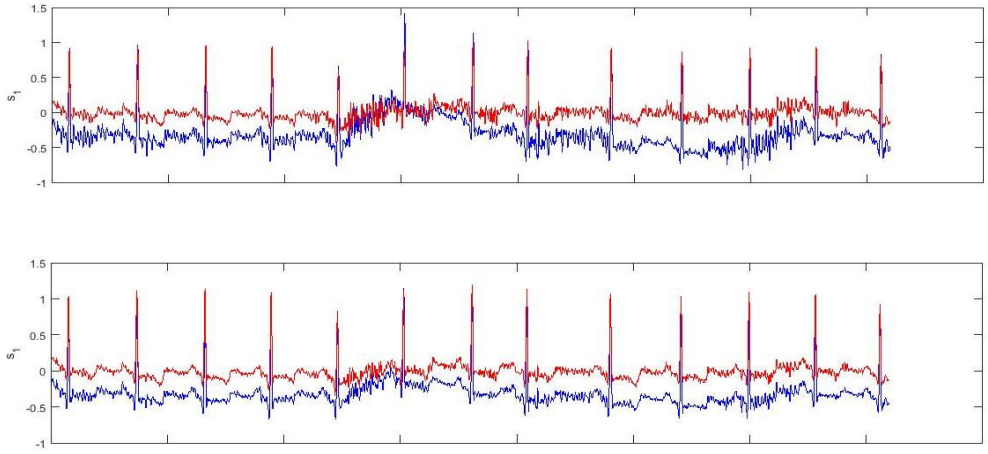

Figure 6.6. Muscle artifact filtering using RobustICA

In Figure 7.6. the top panel the blue line shows ECG signal combined with MA SNR =0 DB and the red line shows the estimated noise free ECG signal. In the bottom panel, the blue line shows ECG signal combined with MA SNR= 6 DB and the red line shows the estimated noise free ECG signal.

Table 6.3. Results of muscle artifact removing using RobustICA.

Muscle artifact (MA)					
SNR DB before	SNR DB after	RMSE	RMSE DB	kurtosis of Ic1	kurtosis of Ic2
-20	-3.5864	9.3013%	-10.3146	3.788	1.3629
-12	4.4136	9.0599%	-10.4288	-1.6328	-1.3626
-6	10.4136	8.6764%	-10.6166	7.6336	2.6656
0	16.4136	8.2643%	-10.8279	16.8079	2.5662
6	22.4136	7.9768%	-10.9817	22.5893	3.7377
12	28.4136	7.5964%	-11.1939	25.6348	7.7858
20	36.4136	7.5412%	-11.2256	31.1994	15.5925

From the above tables, we can conclude that the signal to noise ratio improvement after using RoboustICA in case of baseline wander reduction is about +12 dB, in electrode movement reduction is about +13.5 dB, and in muscle artifact removal is about +16.5 dB.

6.5. Results of Applying Extended Kalman Filter

With the intention of calculating, the effectiveness of EKF in ECG noises reduction a file 100 from MIT-BIH Arrhythmia database has been selected and combined with baseline wandering, EMG artifact, and electrode motion noises taken from MIT-BIH non-stress test database. The selected files has been segmented and mixed to produce a noisy signal have signal to noise ratio values between 20 to -20 dB. The following tables and figures show the signal to noise ratio before and after using RobustICA algorithm. The following tables show the signal to noise ratio before and after using Extended Kalman Filter.

6.5.1. Baseline wander filtering using EKF

Baseline wandering noise has been mixed with a normal ECG signal to obtain noisy signals with signal to noise ratio (20, 12, 6, 0, -6, -12, -20) dB. Table 7.4. contains the values of SNR for noisy ECG signal and the values of SNR for the estimated filtered signals after using EKF for signal filtering.

Table 6.4. Results of baseline wander removing using EKF.

Baseline wander (BW)			
SNR DB before	SNR DB after EKF	SNR DB after EKS	Model Error
-20	-17.9127	-17.9128	0.069348
-12	-9.9128	-9.9227	0.09307
-6	-3.9756	-3.9749	0.25082
0	1.9599	1.9616	0.18924
6	7.2179	7.1773	0.1081
12	11.4462	11.5209	0.37787
20	21.8105	21.8869	0.14487

6.5.2. Electrode motion noise filtering using EKF

Electrode motion noise has been mixed with a normal ECG signal to obtain noisy signals with signal to noise ratio (20, 12, 6, 0, -6, -12, -20) dB. Table 7.5. contains the values of SNR for noisy ECG signal and the values of SNR for the estimated filtered signals after using EKF for signal filtering.

Table 6.5. Results of electrode motion removing using EKF.

Electrode movement (EM)			
SNR DB before	SNR DB after EKF	SNR DB after EKS	Model Error
-20	-18.5713	-18.529	0.18075
-12	-10.9133	-10.9098	0.10398
-6	-4.9318	-4.9077	0.12401
0	1.413	1.5004	0.19329
6	7.7827	7.8614	0.15964
12	13.0413	13.003	0.14153
20	18.4107	18.55	0.13188

6.5.3. Muscle artifact filtering using EKF

Muscle contraction artifact has been mixed with a normal ECG signal to obtain noisy signals with signal to noise ratio (20, 12, 6, 0, -6, -12, -20) dB. Table 7.6. contains the values of SNR for noisy ECG signal and the values of SNR for the estimated filtered signals after using EKF for signal filtering.

Table 6.6. Results of muscle artifact removing using EKF.

Muscle artifact (MA)			
SNR DB before	SNR DB after EKF	SNR DB after EKS	Model Error
-20	-12.7159	-12.0188	0.93915
-12	-4.7445	-4.0378	0.57672
-6	1.0498	1.5993	0.26324
0	7.4387	7.872	0.24274
6	13.1576	13.6908	0.14494
12	18.1413	18.9238	0.15176
20	20.7092	21.5762	0.13565

From the above tables, we can conclude that the signal to noise ratio improvement after using EKF in case of baseline wander reduction is about +2 DB, in electrode movement reduction is about +1.5 DB, and in muscle artifact removal is about +7 DB in the signals which have high signal to noise ratio and it reduced when the noise ratio decrease.

CHAPTER 7. CONCLUSIONS

This research has been concerned with studying the ECG signal and the noises which mask it in order to develop an efficient and reliable algorithm to reduce muscle contraction artifact, baseline shifting, and electrode motion noises as the most common ECG noises. Two algorithms have been applied, one is the source separation method which separates the contaminated observed signal into independent components and second is a model based filtering using Extended Kalman Filter.

The proposed RobustICA algorithm is connected with a correlation based method to estimate a clean ECG signal from the separated independent. While the Extended Kalman Filter using the dynamic nonlinear state-space model of ECG signal to estimate a clean ECG signal. With the intention of evaluating the success of the recommended method an artificially generated database has been created by mixing noise-free ECG signals with baseline wander, muscle artifact, and electrode movement noises to produce signals with SNR 20, 12, 6, 0,-6, -12, and -20 DB.

The results of the analysis illustrate that RobustICA produces better results in the reduction of muscle artifact when compared with baseline wander and electrode movement artifacts reduction ratio, while EKF shows good results in the reduction of muscle artifact when the signals have low SNR. The analysis results also demonstrate that RobustICA is better than EKF in the reduction of baseline wander and electrode movement artifacts while both show good result in the reduction of muscle artifact.

Finally, we can say that RobustICA based estimation algorithm performed better than the EKF method which has produced better SNR in the presence of all common types of artifacts.

REFERENCES

[1] U. Rajendra Acharya, J. S. Suri, J. A. E. Spaan, and S. M. Krishnan, *Advances in cardiac signal processing*. 2007.

[2] G. D. Clifford, F. Azuaje, and P. E. McSharry, *Advanced Methods and Tools for ECG Data Analysis*, vol. 91. 2005.

[3] H. Limaye and V. V Deshmukh, "ECG Noise Sources and Various Noise Removal Techniques : A Survey," vol. 5, no. 2, pp. 86–92, 2016.

[4] P. Papoulis, "Probability, Random Variables, and Stochastic Processes." New York: McGraw-Hill, 1991.

[5] A. J. Flewelling, J. Currie, C. A. Gray, and J. A. Johnson, "Adaptive Wavelet Wiener Filtering of ECG Signals," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 2, pp. 437–445, 2013.

[6] A. Prof, "ECG Signal Filtering using an Improved Wavelet Wiener Filtering," vol. 07, no. 07, pp. 1242–1247, 2015.

[7] A. Hyvärinen, J. Hurri, and P. O. Hoyer, "Independent Component Analysis" pp. 151–175, 2009.

[8] Z. Barati and A. Ayatollahi, "Baseline wandering removal by using independent component analysis to single-channel ECG data," *ICBPE 2006 - Proc. 2006 Int. Conf. Biomed. Pharm. Eng.*, no. 1, pp. 152–156, 2006.

[9] P. M. Mrinal Phegade and U. S. Sutar, "Hybrid ICA Algorithm for ECG Analysis," in *2012 12th International Conference on Hybrid Intelligent Systems (HIS)*, 2012, vol. 67, pp. 478–483.

[10] A. S. Barhatte, R. Ghongade, and S. V. Tekale, "Noise analysis of ECG signal using fast ICA," *Conf. Adv. Signal Process. CASP 2016*, pp. 118–122, 2016.

[11] R. Sameni, M. B. Shamsollahi, C. Jutten, and M. Babaie-Zadeh, "Filtering noisy ECG signals using the extended Kalman filter based on a modified dynamic ECG model," *Comput. Cardiol.*, vol. 32, no. October, pp. 1017–1020, 2005.

[12] G. D. Clifford, A. Shoeb, P. E. McSharry, and B. A. Janz, "Model-based filtering, compression and classification of the ECG," *Int. J. Bioelectromagn.*, vol. 7, no. 1, pp. 158–161, 2005.

[13] R. Sameni, M. B. Shamsollahi, and C. Jutten, “Model-based Bayesian filtering of cardiac contaminants from biomedical recordings,” *Physiol. Meas.*, vol. 29, no. 5, pp. 595–613, 2008.

[14] O. Sayadi, R. Sameni, and M. B. Shamsollahi, “ECG denoising using parameters of ECG dynamical model as the states of an extended kalman filter,” *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, no. September, pp. 2548–2551, 2007.

[15] O. Sayadi, S. Member, and M. B. Shamsollahi, “ECG Denoising and Compression Using a Modified Extended Kalman Filter Structure,” vol. 55, no. 9, pp. 2240–2248, 2008.

[16] O. Sayadi, M. B. Shamsollahi, and G. D. Clifford, “Synthetic ECG generation and Bayesian filtering using a Gaussian wave-based dynamical model,” *Physiol. Meas.*, vol. 31, no. 10, pp. 1309–1329, 2010.

[17] R. Vullings, B. De Vries, and J. W. M. Bergmans, “An adaptive Kalman filter for ECG signal enhancement,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 4, pp. 1094–1103, 2011.

[18] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, “A Dynamical Model for Generating Synthetic Electrocardiogram Signals,” *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289–294, 2003.

[19] M. Sarfraz, “Role of Independent Component Analysis in Intelligent Ecg Signal Processing,” *Sch. Comput. , Sci. Eng. Univ. Salford , Salford , UK*, no. August, 2014.

[20] V. Zarzoso and P. Comon, “Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size,” *IEEE Trans. Neural Networks*, vol. 21, no. 2, pp. 248–261, 2010.

[21] P. C. Vicente Zarzoso and M. Kallel, “HOW FAST IS FASTICA ?,” in *In Proc. EUSIPCO-2006*, 2006.

[22] R. Sameni, M. B. Shamsollahi, and C. Jutten, “Filtering Electrocardiogram Signals Using the Extended Kalman Filter,” in *Engineering in Medicine and Biology*, 2006, pp. 5639–5642.

RESUME

Tasnim Ahmed Abdelrazig Mohammed, I was born on 20.08.1992 in Jeddah. I finished my primary and secondery school in Wad Medani, Sudan. I started my Bachlor at Gezira university, Faculty of Engineering and Technology in 2009. I graduated from Deparment of Electronics Engineering, Branch of medical instrumentation in 2014. I came to Turkey in 2015 and started learning Turkish, then I started my Master at Sakarya university in 2016.