

**ROLE OF INDEPENDENT COMPONENT
ANALYSIS IN INTELLIGENT ECG SIGNAL
PROCESSING**

Mohammad Sarfraz

School of Computing, Science and Engineering

University of Salford, Salford, UK

Submitted in Partial Fulfilment of the Requirement of
the Degree of Doctor of Philosophy, August 2014

Contents

Contents	i
List of tables	iv
List of figures	v
Acknowledgements	ix
Abbreviations	x
Abstract	xii
Chapter 1 Introduction	1
Chapter 2 Background & Literature Review	4
Chapter 3 Heart and ECG	14
3.1 Heart Anatomy	14
3.2 Sino Atrial (S-A) Node.....	16
3.3 Depolarisation	17
3.4 Generation of Ionic Current.....	19
3.5 Electrocardiogram	20
3.6 ECG Wave Description	21
3.7 ECG Lead Systems.....	23
3.8 Ambulatory ECG.....	28
3.9 Heart Rhythms and Arrhythmias.....	29

3.9.1 Right Bundle Branch Block (RBBB).....	29
3.9.2 Left Bundle Branch Block (LBBB)	30
3.9.3 Premature Ventricular Contraction (PVC).....	31
3.9.4 Paced Beat.....	31
3.9.5 Atrial Premature Beat.....	32
3.9.6 Flutter Wave.....	32
3.9.7 Ventricular Escape Beat.....	32
3.10 Noise in ECG.....	33
Chapter 4 Independent Component Analysis.....	36
4.1 Independence.....	36
4.2 Illustrative Example of ICA	38
4.3 Probability Density Function.....	43
4.4 Central Limit Theorem.....	44
4.5 Moments.....	45
4.6 Preprocessing.....	48
4.6.1 Centering.....	49
4.6.2 Whitening.....	49
4.7 ICA Algorithms	51
4.7.1 Infomax ICA	51

4.7.2 FastICA Algorithm	54
4.7.3 JADE Algorithm	56
Chapter 5 Principal Component Analysis	58
Chapter 6 Artificial Neural Network.....	64
6.1 Network Performance Analysis.....	65
6.2 Back Propagation Network.....	66
6.3 Support Vector Machine	70
Chapter 7 Proposed Method- Hypothesis	73
7.1 Motion Artefacts Removal: Proposed Method.....	73
7.2 Proposed Algorithm	73
7.3 Classification with Feature extraction from ECG using	80
7.4 Proposed Method.....	81
Chapter 8 Results and Discussion	89
8.1 Discussion	114
Chapter 9 Conclusion and Future Work	116
9.1 Propose and develop the algorithm for denoising of ECG signals using ICA	116
9.2 Proposed and developed the algorithm for arrhythmia classification using ICA.....	117
References	120
Appendix: Related Publication	135

List of tables

Table 7-1 Record and number of ECG sample used.....	85
Table 8-1 Comparison of classification Sensitivity before and after source separation.	94
Table 8-2 Comparison of classification positive predictive value before and after source separation.	95
Table 8-3 Comparison of classification accuracy before and after source separation.	96
Table 8-4 Detailed accuracy for different arrhythmias with proposed method	104
Table 8-5 Detailed positive predictive value for different arrhythmias with proposed method	105
Table 8-6 Detailed sensitivity for different arrhythmias with the proposed method	106
Table 8-7 Detailed accuracy for different arrhythmias with SVM classifier	108
Table 8-8 Detailed positive predictive value for different arrhythmias with SVM classifier	109
Table 8-9 Detailed sensitivity for different arrhythmias with SVM classifier.....	110
Table 8-10 Detailed specificity for different arrhythmias with PCA extracted features	112
Table 8-11 Detailed sensitivity for different arrhythmias with PCA extracted features.....	113
Table 8-12 Detailed accuracy for different arrhythmias with PCA extracted features.....	113

List of figures

Figure 3-1 Anatomy of the heart and associated vessels.	15
Figure 3-2 Electrophysiology of the cardiac muscle cell.....	18
Figure 3-3 Basic ECG wave shape with morphological features	21
Figure 3-4 Einthoven triangle and ECG limb leads definition	23
Figure 3-5 Augmented lead placement	25
Figure 3-6 six chest electrodes placement for ECG.....	26
Figure 3-7 Different component of ECG during depolarization of atria and ventricle.....	27
Figure 3-8 RBBB arrhythmia recorded by different leads.....	30
Figure 3-9 Premature Ventricular Contraction	31
Figure 4-1 ICA demonstration with Cocktail Party Problem	37
Figure 4-2 Blind source separation (BSS) block diagram	38
Figure 4-3 ICA demonstration original source signal.....	39
Figure 4-4 ICA demonstration mixed (observed) signal.....	40
Figure 4-5 ICA demonstration recovered source signal	41
Figure 4-6 Whitening Process.....	50
Figure 4-7 Learning objectives of Infomax ICA algorithm.....	52
Figure 4-8 Mixing and de-mixing model.....	53
Figure 5-1 Demonstration of difference between PCA and ICA.....	60

Figure 6-1 Layout of ANN.....	65
Figure 6-2 Three layer back-propagation neural network.....	66
Figure 6-3 SVM methodology of maximising margin between different classes	70
Figure 7-1 Proposed setup for removing motion artefacts.....	75
Figure 7-2 Motion artefact removal flowchart.....	76
Figure 7-3 Sample of ‘em’ noise 10 Sec.....	77
Figure 7-4 Ambulatory ECG with motion artefact	77
Figure 7-5 Extracted signal from two leads	78
Figure 7-6 Original ECG data.....	78
Figure 7-7 A normal heartbeat is a linear combination of ICA bases with Coefficients,	82
Figure 7-8 ECG heartbeats of eight types	83
Figure 7-9 Block diagram of the proposed feature extraction and classification system	86
Figure 8-1 ECG Combine with noise (Top Panel), ‘em’ Noise middle panel, Extracted ECG with ICA bottom panel. Y Axis is normalized after ICA	92
Figure 8-2 Pan Tomkins algorithm for beat detection on a sample of a clean ECG, a noisy ECG (SNR= -6 dB) & filtered ECG using ICA. The beat detection has significant improvement after ICA filtering.	93
Figure 8-3 Sensitivity with ICA filtering and without filtering	95
Figure 8-4 Positive Predictivity with ICA filtering and without filtering.....	96
Figure 8-5 Accuracy with ICA filtering and without filtering.....	97

Figure 8-6 ECG beat detection using a Pan Tomkins algorithm	97
Figure 8-7 ECG QRS detection using a Pan Tomkins algorithm	98
Figure 8-8 Beat detection comparison for LBBB arrhythmia with proposed ICA filtering and without filtering.....	98
Figure 8-9 Beat detection comparison for Normal ECG beat with proposed ICA filtering and without filtering.....	99
Figure 8-10 Beat detection comparison for RBBB ECG beat with proposed ICA filtering and without filtering.....	99
Figure 8-11 Beat detection comparison for VEB arrhythmia with proposed ICA filtering and without filtering.....	100
Figure 8-12 Classification into eight classes with proposed method.....	103
Figure 8-13 Results of Arrhythmia Classification Accuracy with proposed method	104
Figure 8-14 Results of positive predictive value with the proposed method.....	105
Figure 8-15 Results of sensitivity with the proposed method.....	106
Figure 8-16 Results of Arrhythmia Classification Accuracy with SVM classifier.....	107
Figure 8-17 Results of positive predictive value with SVM classifier	108
Figure 8-18 Results of sensitivity with SVM classifier	109
Figure 8-19 Results of accuracy with PCA extracted features	111
Figure 8-20 Results of positive predictive value with PCA extracted features	111
Figure 8-21 Results of positive predictive value with PCA extracted features	112

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr Francis Li for the continuous support of my Ph.D. study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me throughout my research and writing of this thesis. Without his invaluable support and encouragement, this thesis would not have been accomplished. He has contributed greatly to every paper I have written, and to every progress I have made.

I would also like to extend my gratitude's to Professor P Gaydecki and Professor T Rich ling for their keen interest in my thesis. Their observations are a great source of knowledge as well as realization of some weak areas in my work. I hope their observations will go a long way in my further research endeavours.

My sincere thanks also go to Dr Ateeq Ahmed Khan for being my local supervisor and offering me every support at Salman bin Abdul Aziz University. He has given much help and many suggestions on not only my research and study, but also on my life in Alkharj. I would also like to thank my Dean Dr Osama Saad Alqahtani, College of Engineering, Salman bin Abdul-Aziz University, for his support and motivation for providing me all the help at Salman bin Abdul Aziz University to complete this thesis.

Last but not the least, I would like to thank my family, parents and my wife Mariyam Siddique for her understanding and love, she has walked everyday with me, showed endless grace. I would have never completed this without her.

Abbreviations

A	Atrial Premature Beat
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
ANSI	American National Standards Institute
AV	Atrio Ventricular
BPNN	Back Propagation Neural Network
CVD	Cardiovascular Diseases
ECG	Electrocardiogram
EEG	Electroencephalogram
ERP	Event-Related Potentials
F	Fusion Of Paced And Normal Beats
GUI	Graphical User Interface
HBR	Heart Beat Rate
ICA	Independent Component Analysis
ISO	Isoelectric Line
LMS	Least Mean Square
LVQ	Linear Vector Quantization
MART	Multi-Channel Adaptive Resonance Theory
MIT-BIH	Massachusetts Institute Of Technology Beth
MLP	Multi Layer Perceptron
MSE	Mean Squared Error
N	Normal
NLMS	Normalized LMS Algorithms
P	Paced Beats
PCA	Principal Component Analysis
PPV	Positive Predictive Value
PVC	Premature Ventricular Contraction
QRS	Combination Of 3 Graphical Defection On

R	Right Bundle Branch Block
RBBB	Right Bundle Branch Block
RCE	Restricted Coulomb Energy
SA	Sino Atrial
SNR	Signal To Noise Ratio
SOM	Self-Organization Map
STD	Standard Deviation
em	Electrode Motion Artefact
WHO	World Health Organisation

Abstract

The Electrocardiogram (ECG) reflects the activities and the attributes of the human heart and reveals very important hidden information in its structure. The information is extracted by means of ECG signal analysis to gain insights that are very crucial in explaining and identifying various pathological conditions. The feature extraction process can be accomplished directly by an expert through, visual inspection of ECGs printed on paper or displayed on a screen. However, the complexity and the time taken for the ECG signals to be visually inspected and manually analysed means that it's a very tedious task thus yielding limited descriptions. In addition, a manual ECG analysis is always prone to errors: human oversights.

Moreover ECG signal processing has become a prevalent and effective tool for research and clinical practices. A typical computer based ECG analysis system includes a signal pre-processing, beats detection and feature extraction stages, followed by classification. Automatic identification of arrhythmias from the ECG is one important biomedical application of pattern recognition. This thesis focuses on ECG signal processing using Independent Component Analysis (ICA), which has received increasing attention as a signal conditioning and feature extraction technique for biomedical application.

Long term ECG monitoring is often required to reliably identify the arrhythmia. Motion induced artefacts are particularly common in ambulatory and Holter recordings, which are difficult to remove with conventional filters due to their similarity to the shape of ectopic

beats. Feature selection has always been an important step towards more accurate, reliable and speedy pattern recognition. Better feature spaces are also sought after in ECG pattern recognition applications.

Two new algorithms are proposed, developed and validated in this thesis, one for removing non-trivial noises in ECGs using the ICA and the other deploys the ICA extracted features to improve recognition of arrhythmias. Firstly, independent component analysis has been studied and found effective in this PhD project to separate out motion induced artefacts in ECGs, the independent component corresponding to noise is then removed from the ECG according to kurtosis and correlation measurement.

The second algorithm has been developed for ECG feature extraction, in which the independent component analysis has been used to obtain a set of features, or basis functions of the ECG signals generated hypothetically by different parts of the heart during the normal and arrhythmic cardiac cycle. ECGs are then classified based on the basis functions along with other time domain features. The selection of the appropriate feature set for classifier has been found important for better performance and quicker response. Artificial neural networks based pattern recognition engines are used to perform final classification to measure the performance of ICA extracted features and effectiveness of the ICA based artefacts reduction algorithm.

The motion artefacts are effectively removed from the ECG signal which is shown by beat detection on noisy and cleaned ECG signals after ICA processing. Using the ICA extracted feature sets classification of ECG arrhythmia into eight classes with fewer independent components and very high classification accuracy is achieved.

Chapter 1 Introduction

Many components of the measured electrocardiogram (ECG) signal originate from different and hypothetically independent sources, independent sources; the joint effect is a linear combination of them presented at the ECG electrodes. In the ECG signal processing there are several unsolved problems and many existing solutions that need optimisation, noise reduction/removal is one of them. There are several techniques for noise removal from the ECG which can be employed to give good performance results in the controlled environment. But in some cases to correctly identify the arrhythmia, long-term ECG monitoring is required, which is often acquired with ambulatory ECG usually recorded with Holter device, where it is difficult to obtain controlled environment setting. This makes the ECG recording more susceptible to different kind of noises not commonly witnessed at controlled environment. There are several filtering techniques that can be employed to remove some of the noises according to different frequencies, e.g. power interference baseline wandering, etc., but not electrode motion artefacts known in ECG studies as ‘em’ noise which results from the motion of electrode on the patient’s skin or due to the movement of the patient itself. The motion artefact is very difficult to be removed by conventional filters because of its ectopic in nature; it takes the shape of the wave which makes it difficult to be removed. ICA presents the solution for removing the motion artefact enabling further processing of ECG signals. Due to the different origination of the ECG and the motion artefact, the signals are independent of each other ECG having a super Gaussian distribution can be extracted from the noisy signal.

In this work ICA algorithm will be used for ECG signal processing. Firstly, to develop a method to detect, estimate and remove motion artefacts from the ECG. Secondly, to develop an algorithm for ECG arrhythmia classification using features extracted with ICA. A new feature set is developed which includes morphological and dynamic ECG features along with ICA extracted features. This makes a robust feature set for ECG arrhythmia classification. Comparison of classification accuracy, specificity and sensitivity with other state of the art method is done and its efficiency is proven.

More Specifically

The aim of this thesis is to investigate the roles of Independent Component Analysis (ICA) in Electrocardiogram (ECG) signal processing. “Can the ICA be effectively used to mitigate noise from other sources and improve the performance of automatic classification?” is the research question.

In order to accomplish this aim, the following objectives are to be fulfilled:

- Identification and/or preparation of a suitable dataset to carry out an empirical study of the associated signal processing and machine learning algorithms.
- Identification of suitable ECG data representation, alignment methods.
- Identification of suitable ICA algorithms for ECG applications.
- Developing signal probing acquisition technique methods
- Developing a feature extraction set using ICA.
- Developing related signal processing, machine learning algorithms.
- Critically evaluating the method developed in this study and identifying its limitations.

The organisation of the thesis is as follows, Chapter 1 is the general introduction. Chapter 2 presents the background and literature survey, including the current state of art in ECG signal processing techniques. Chapter 3 describes the anatomy of the heart, the generation of electric impulse from Sino atrial nodes and ionic current and presents a summary of ECG techniques including signal detection, description, properties of ECG wave and noises in ECG along with summary of different arrhythmia in the ECG. Chapter 4 describes the independent component analysis and various ICA algorithms. Chapter 5 discusses the principal component analysis and its application in ECG . Chapter 6 studies neural networks with a view of different architectures, algorithms and activation function . Chapter 7 present proposed method & hypothesis for noise removal and classification algorithms. Chapter 8 summarises result of this work followed by a discussion. Finally, Chapter 9 contains a conclusion and future work.

Chapter 2 Background & Literature Review

The occurrence and prevalence of cardiovascular diseases (CVD) have increased in recent years (Anand & Yusuf, 2011). As per the recent WHO report the overall death rate due to CVD in 2007 is 251.7 per 100,00 population. (De et al., 2009). This increase in death rates due to CVD in the modern world is attributed to the increasing number of patients of obesity, diabetes mellitus, smoking habit and other lifestyle changes. One of the complication of CVD among many others is atrial and ventricular arrhythmias, which happen due to cardiac rhythm disturbances. Arrhythmia is a collective term for a heterogeneous group of conditions in which there would be an abnormal electrical activity. There are many causes for arrhythmias, most of which are related to CVD. Arrhythmias like ventricular fibrillation and flutter are life threatening medical emergencies which result in cardiac arrest, hemodynamic collapse and sudden cardiac death. (Huikuri, Castellanos, & Myerburg, 2001).

Electrocardiogram (ECG) has become one of the important tools in the diagnosis of heart disease. Due to high mortality rate early detection and correct identification of ECG arrhythmia is very important. ECG signal processing is started with reduction of noise and artefacts to make ECG signal cleaner which can be interpreted easily by manual inspection or by automatic diagnosis machine. Then important hidden features are extracted which are not easily identified by manual inspection. These features are used for automatic ECG diagnosis, which classifies them into different classes of normal or irrational heartbeat. Electrocardiography has an important role in cardiology since it consists of effective, simple,

non-invasive, low cost way for the diagnosis of cardiac disorders which have a high rate of occurrence and are very relevant to their impact on patient life and social costs (Linh & Osowski, 2003).

Noise Removal from ECG

Noise and artefact removal is the first step for ECG signal processing (Wisbeck, Barros, Yy, & Ojeda, 1998) used ICA for removing breathing artefact with promising results which led them to apply ICA technique for more noise separation . (Barros, Mansour, & Ohnishi, 1998) presented their work by using well established MIT-BIT noise stress database. They proposed ICA based architecture for BSS separation of linearly mixed signals. The architecture consisted of a high-pass filter, a two-layer network based on ICA algorithm and a self-adaptive step-size. The step-size is derived from the mean behaviour of output signals. The two layered algorithm provided fast convergence as compared to other algorithms which used whitening technique along with ICA algorithm.

Independent component analysis can be implemented with different algorithms; each has its own merits, as they can be problem specific. For the case of noise and artefact removal from ECG (Sarfranz et al., 2011) , performed comparative study of different ICA algorithms for ECG signal processing. Some motion artefact are ectopic in nature hence they cannot be easily detected by conventional filters (I. Romero, 2011) used PCA-ICA based algorithms for motion artefact removal. Carrying the idea forward (Sarfranz & Li, 2013) used two lead design for motion artefact removal along with feature extraction of ECG using ICA , which was extension to previous work.

(I. Romero, 2011; Inaki Romero, 2010) proposed the use of PCA and ICA for noise reduction and motion artefact removal, for ambulatory conditions where noise increases relative to activity, the signal quality is greatly reduced by motion artefact that are ectopic in nature. Making it difficult to be removed by conventional filters. ECG is collected from the lumbar curve region where ECG signals are negligible. To design a completely automatic system for diagnosing ECG using PCA and ICA, kurtosis is used to identify signal from noise by putting a threshold value as a classifier. It would be interesting to see how other automatic component selection work in this case.

ECG Feature Extraction

Independent component analysis ((Hyvärinen, Karhunen, & Oja, 2001; Hyvärinen & Oja, 2000a; Naik & Kumar, 2011; Owis, Youssef, & Kadah, 2002) which is a form of blind source separation method is a statistical signal processing technique used for separating a set of signals into mutually independent component signals.

ICA for feature extraction (Huang, Hu, & Zhu, 2012; Hyvärinen et al., 2001; Jiang, Zhang, Zhao, & Albayrak, 2006b; S.-N. Yu & Chou, 2006, 2008) is a new area of research which has good potential for getting important features to be used by classifier for automatic ECG classification. Features which can be used for ECG classification can be categorized according to their selection process. (Afonso, Tompkins, Nguyen, & Luo, 1999; De Chazal, O'Dwyer, & Reilly, 2004) used time domain methods, (R. Acharya et al., 2004; Al-Fahoum & Howitt, 1999) used transformation methods for feature extraction. (Kwak, Choi, & Choi, 2001; Kwak & Choi, 2003) proposed a feature selection algorithm where output class information was included with input features; the added information about the class improved

the performance in the extraction of useful features for classification. Standard feed-forward neural network was used for classification stage, results are promising, but it fell short of applying this to dataset other than bio signals and multiclass issue was also not taken into consideration.

(Owis et al., 2002) illustrates ICA as feature extractor for ECG. ICA was used to calculate independent components (ICs) of the selected ECG signal, ICs are then used as bases for extraction of important features, 219 ICs are used and 100% classification accuracy is achieved in normal beats classification and moderate accuracy is achieved in another four arrhythmia classification. Additionally computational time was also high and even more for larger datasets. Carrying the same idea (Jiang et al., 2006b) introduced ICA feature extraction method joining ICA basis function coefficient and wavelet transform coefficient forming an over complete feature vector. Mutual information feature selection algorithm(Hyvärinen & Oja, 2000b) is used to select important features. Classification is done by support vector machine algorithms varied accuracy for different beats is achieved 77 % for atrial flutter to 98% for normal beats.

(S.-N. Yu & Chou, 2007) applied ICA for feature extraction in time domain using minimum distance and bayes classifiers for classification of six different type of ECG beats. Switchable scheme was proposed using two ICA based features of different lengths. The selection between two algorithms is made using RR interval as an indicator. High accuracy was achieved with this approach. However only simple classifiers were studied and more than six types of beat classification were not covered. In the following year (S.-N. Yu & Chou, 2008) assessed the integration of independent component analysis and neural network as classifier

for eight ECG beats classification. Only one bank of ICA classifier was used in time domain , which works as basic features for ECG signal representation. Two neural networks models a probabilistic neural network and a back propagation network were studied and results were significant. Continuing their work further (S.-N. Yu & Chou, 2009) proposed novel ICs arrangement strategy based on the L2 norms of the rows of Demixing Matrix. To reduce the computational burden on the system and processing time, only important ICs features are selected to be part of feature vector. This process makes smaller feature set while maintaining the desired accuracy level. Results were promising and further application of this algorithm on other data sets and different noise and artefact must be done for deeper analysis.

(Y. Wu & Zhang, 2011) used redundancy and relevance optimization for feature selection criteria and support vector machine for classification with 14 heart beat types achieving stable accuracy of 90 % in all the cases as compared to (Jiang et al., 2006b) work where high accuracy is achieved in some cases. (Shen, Wang, Zhu, & Zhu, 2010) also used Independent component analysis and support vector machine for multi lead ECG classification. Unlike others they segmented heartbeat into three segments, P wave , QRS interval and ST segment) separate features for each segment are extracted using ICA then combined to form classifier for 11 heart beat types. Structure of training data and amount of data is identified as possible area for performance improvement.

As claimed by some researchers that ICA performs better with more leads for better understanding of ECG features. (Zhu et al., 2008) examined 98-lead and 72-lead ECG data and by using ICA they were able to recover separate ECG feature of P wave, QRS and T wave. Thus high resolution data can be used in detailed diagnosis.

However, all these methods have some drawbacks. Most of these methods were tested only on limited data sets and the generalization performance of these methods on large databases was not tested. All these methods are tested only on a few classes of ECG beats. There is a need to test the methods and algorithms on a standard classification scheme of arrhythmia beats such as ANSI/AAMI EC57:1998 (EC57, 1998; Martis, Acharya, & Min, 2013)

A number of practical applications of ICA in source separation are mentioned in the literature. To mention some, ICA are used in finance industry for financial data feature extraction and forecasting of time series. (Cheung & Xu, 2001; Lu, Lee, & Chiu, 2009), (Bingham, Kuusisto, & Lagus, 2002; Kolenda, Hansen, & Sigurdsson, 2000; Pu & Yang, 2006) used ICA for analysis of document text.(Acerese et al., 2004; Cabras, Carniel, & Isserman, 2010) studied application of ICA in seismic monitoring for volcanic activity analysis and prediction.

ECG Classification

Neural nets are most suited for situations that bear no clear set of rules or relationships. Because neural nets learn from experience, they can do tasks that are otherwise hard to identify precisely. Neural nets can classify complex data. They are applied successfully to a wide range of application like biomedical signal analysis, financial time series forecasting and process control (Rosaria & Carlo, 1998)(Mark & Jansen, 1997) (Fikret, 1999; Hu, Xie, & Tan, 2004; S. Wu, Chow, & Using, 2004).

(Guvendir, Acar, Demiroz, & Cekin, 1997) used Neural Networks for ECG classification problems. The Arrhythmia dataset of UCI (Bache & Lichman, 2013) was used The dataset

contained 279 traits, 206 of which were linear-valued and the rest were nominal. There were 245 examples of normal ECG and 207 of 15 types of arrhythmia, including coronary artery disease, myocardial infarction, sinus tachycardia, sinus bradycardia, right bundle branch block, and atrial fibrillation. In order to treat the problem as a binary classification one, the 207 examples of arrhythmia are combined in a single class.

(Barro, Fernández-Delgado, Vila-Sobrino, Regueiro, & Sánchez, 1998) adopted multichannel adaptive resonance theory (MART) to classify the ECG patterns. Application results show that this classifier can classify normal and ventricular beats with an accuracy of more than 90%. (Olmez, 1997) worked on classification of ECG waveforms using Restricted Coulomb Energy (RCE), neural networks and genetic algorithms. He detected 4 types of beat such as normal beat, left bundle branch block beat, premature ventricular contraction and paced beat. He obtained a classification accuracy of more than 94%.

(Dokur & Ölmez, 2001) used multi-layer perceptron to classify ten arrhythmias with training accuracy of 78% using discrete fourier transform for feature extraction. This work is followed by (Güler & Übeyli, 2005; Güler, 2005) for ECG classification using discrete wavelet transform along with multilayer perceptron getting 97.98% training accuracy with five classes of arrhythmia.(Asl, Setarehdan, & Mohebbi, 2008a; Mohammadzadeh-Asl & Setarehdan, 2006)(Asl, Setarehdan, & Mohebbi, 2008b) produced series of work for ECG classification using artificial neural network along with varied feature reduction techniques. Six classes of arrhythmia were identified; training accuracy of 98.22% was achieved using only multilayer perceptron. This work was followed by introduction of feature reduction technique of

principal component analysis and linear discriminant analysis in addition to MLP for ECG classification.

Neural networks' origin can be tracked down to the domain of cognitive engineering, it is thought to be suitable to do complex jobs. They are particularly efficient at spotting patterns, sorting information, and processing noisy signals. Neural networks are loosely modelled on the networks of neurons in biological systems. Artificial neural networks (Haykin, 1995)(Negnevitsky, 2005) are collections of mathematical models that emulate some of the observed properties of biological nervous systems and draw on the analogies of adaptive biological learning. An artificial neural net (ANN) is an information-processing paradigm inspired by the way the densely interconnected, parallel structure of the brain processes data. The key factor of the ANN model is the novel structure of the data processing system. It is composed of a large number of highly interconnected processing elements that are analogous to neurons and are bound together with weighted connections that are analogous to synapses.

(Nazmy, El-Messiry, & Al-Bokhity, 2010) presents an intelligent diagnosis system using hybrid approach of adaptive neuro-fuzzy inference system (ANFIS) model for classification of Electrocardiogram (ECG) signals. Six types of ECG signals are classified. The results indicate a high level of efficiency of the tools used with an accuracy level of more than 97%. Five power spectrum features were extracted from the ECG signal at (4, 8, 12, 16 and 20 Hz). The term power spectrum means the amount of power per unit (density) of frequency (spectral) as a function of the frequency (Minami, Nakajima, & Toyoshima, 1999).

Other Bio Signals Application

Event-Related Potentials (ERPs) are time series of voltages in the on-going electroencephalogram (EEG) that are time and phase-locked into a set of similar experimental event. However due to overlapping of time courses and scalp projections ERPs cannot be easily decomposed into functionally distinct components (Jung, Makeig, Lee, et al., 2000). ICA is used effectively to decompose multiple overlapping components from sets of related ERPs (Makeig, Bell, Jung, Sejnowski, & others, 1996) (Jung et al., 1998; Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997) using Infomax algorithm for detection of ERP.

(Tichavský, Zima, & Krajca, 2011) used ICA for artefact removal from EEG data of arbitrary length that have relatively short duration and exceed in the magnitude of the neighbourhood signal for examples eye blinks or occasional body movement artefacts.

(Jung, Makeig, Westerfield, et al., 2000) used linear ICA to separate neural activity from muscle and blink artefacts in spontaneous EEG data. This work is followed by another study verifying that the ICA can separate artefactual, stimulus locked, response-locked, and non-event related background EEG activities into separate components (Jung et al., 2001).

(Mishra & Singla, 2013) proposed a technique for the removal of eye blink artefact from EEG and ECG signal using fixed point or FastICA algorithm of Independent Component Analysis as a pre-processor for biometric recognition claiming that every individual has distinct EEG and ECG spectrum.

Surface Electromyography (sEMG) is the electrical recording of the muscle activity from the surface. It is linked to the strength of muscle contraction and is a useful to estimate the strength of contraction of the muscle. But the occurrence of artefacts, especially at low level of muscle activity make the recordings unreliable (Naik, Kumar, & Arjunan, 2010).

Conclusion

Ever since the invention of ECG by Willem Einthoven in 1903, it has been used as an important tool in diagnosis of heart. Noise removal from the ECG was the main theme of researchers working in this area as mentioned in the above studies. With the advent of modern computer technology automatic classification and long term ECG monitoring was the focus. Independent component analysis was effectively used for noise removal from ECG. But they did not address the removal of motion artefacts which take the shape of ectopic beat and cannot be easily removed with conventional filter. Also the use of ICA as feature extraction tool was not fully explored to reduce the computational burden and processing time for automatic classificati of ECG by selection of important features only.

Chapter 3 Heart and ECG

The purpose of this chapter is to present basic facts about heart structure, its function and the basics of ECG measurement and to review the most common lead system used in electrocardiography. This will help in understanding ICA applications on ECG in a better way. If the reader is familiar with heart and ECG they can skip this chapter.

The heart is an efficient muscular organ that pumps blood throughout the whole body. It delivers oxygenated blood and other nutrients to body organs and take away metabolic waste from body organs back to the kidney and lungs respectively for excretion. The heart is a muscular pumping organ responsible for delivering oxygenated blood and other nutrients to body organs and removing deoxygenated blood from body organs to the lungs where the blood is again oxygenated.

3.1 Heart Anatomy

The heart consists of four chambers, the top two chambers are called atria and bottom chambers are the ventricles. Atria's are electrically isolated from the ventricles by a nonconductive fibrous tissue (Rajendra Acharya, 2007). The figure 3-1 shows a cross section view of heart, the four chambers and separating muscular tissue are clearly visible. The atria are separated with each other by a thin interatrial septum, while ventricles are separated from each other by a thick interventricular septum to get more pressure for pumping blood. The heart can be considered as double pump where the right side of the heart consisting of right atria receive deoxygenated blood delivered by veins called as superior and inferior venacava. It then pushes this blood to the right ventricle through tricuspid valve which checks the flow of blood in one direction only. Then the contraction of right ventricle pumps the blood into

the lungs where the blood is oxygenated. Likewise, the left side of heart consisting of left atria and left ventricle that also form a pump. Oxygenated blood from the lungs is received by the left atria through the pulmonary vein. It is then pushed into the left ventricle through the mitral valve. The contraction of left ventricle pushes the blood to other body organs. Blood flow in heart is regulated by valves which prevents the back flow of the blood. ECG can be used to identify faulty valve (Davie et al., 1996).

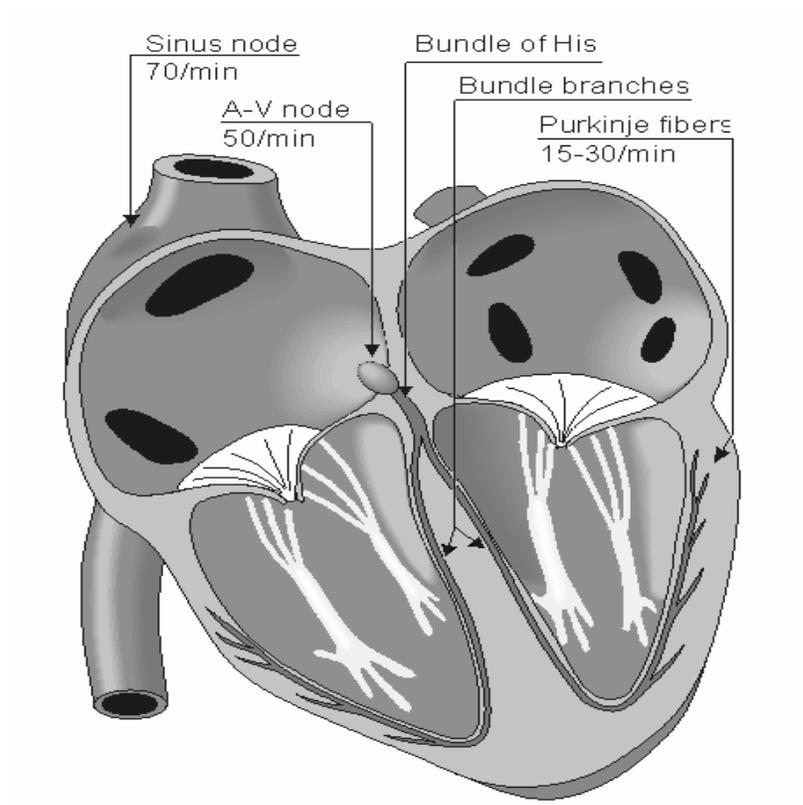


Figure 3-1 Anatomy of the heart and associated vessels.

(Figure Adapted from Malmivuo & Plonsey, 1995)

This pumping process of heart is synchronized. Atria contract first to fill ventricle, then ventricle contract together to pump the blood into the lungs and other parts of the body. The period during which heart is filled with blood is called diastole, and contraction period of

ventricle is called as systole. In electrical terms these two phases are called as repolarization and depolarization. The heart muscles are made up of three layers with inner one as endocardium, middle layer as myocardium, outer layer is called as epicardium. In the right atrium there is an S-A node which is a network of pacemaker cells consisting of excitable tissue that initiates the conduction which travels down the pathways.

Conduction system of surface electrode is used to detect the depolarization of excitable myocardium. When the depolarization propagates towards the positive electrode of the amplifier, the voltage detected is seen as positive and is represented by an upward deflection in the ECG.

The heart can effectively pump the blood to the body organs only when the contraction of different chambers is synchronized. The atria must fill first and then the blood is pumped into the ventricle before the contraction of ventricle pumps the blood to the lungs and other parts of the body. This synchronisation is accomplished by an intricate electrical conduction system that contains the precise timing for the depolarizing of electrically excitable myocardium. This precise control starts with an intrinsic self-excited cardiac pacemaker, which controls the rate at which the heart beats. The pacemaker continuously generates regular electrical impulses, which then go through the conduction system of the heart and initiate contraction of the myocardium. This pacemaker is called the Sino Atrial node or (SA) node.

3.2 Sino Atrial (S-A) Node

The S-A node is located in the upper wall of the right atrium. It is the initial source of electrical excitation under normal conditions. The S-A node is a grid of pacemaker cells

which exhibit automaticity. Automaticity is a property of the cell to get depolarise without any external stimulus. The S-A node sends the depolarization wave to the nearby muscle cells. There are specialised band of tissues between right atrium and left atria known as Bachman's bundles, which provides a fast gateway for the depolarising signal from the right atrium. Due to this fast signal transfer the depolarisation of the right and left atrium occurs almost simultaneously. Automaticity of the S-A node can be affected by sympathetic and parasympathetic inputs or by medicines. There is another special type of tissue, which carries the depolarisation signal from the S-A node to AV node which is called as internodal tracts. The AV node is the starting point of the ventricular conduction system.

3.3 Depolarisation

Depolarisation is a change from negative membrane potential to positive membrane potential. Depolarisation allows the propagation of electrical signals through the cells. There is a potential difference across the cell membrane, due to different concentration of ion's particular Na^+ , K^+ , Cl^- and Ca^{2+} internal and external cellular fluids. When an electrical impulse exceeding the threshold voltage arrives at the cell, the wall of cell membrane becomes permeable allowing the exchange of ions. Once a cardiac step is depolarised, the electrical stimulation is propagated to all the adjacent cells. At the end of depolarisation, the cell membrane again returns to the resting stage and exchange of ions is blocked. All the unwanted ions are pumped out through ion channel in the cell maintaining the ionic balance with the resting state, this process is called repolarization. This repolarisation process is described in figure 3-2 showing electrophysiology process of cardiac muscle cell.

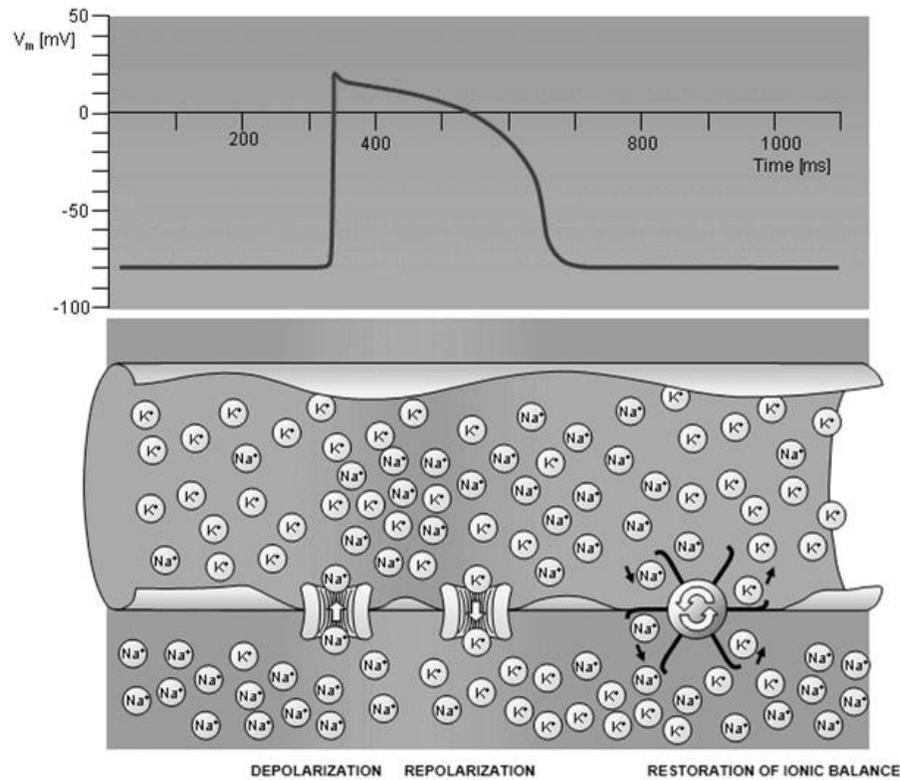


Figure 3-2 Electrophysiology of the cardiac muscle cell
(Figure adapted from Malmivuo & Plonsey, 1995)

The heart muscles are made up of myocardial cells, which are interconnected with each other by the intercalated discs. There are areas on the intercalated discs which are called gap junctions. The gap junctions have low resistance, thus provide a path for rapid conduction of electrical impulses from one cell to another. Strong ionic current is generated due to depolarisation of heart muscles. This current when flows through the resistive body tissues generates a voltage drop which is detected by the electrodes on the skin. Normally it is about 3 mV peak at the chest. Although there exists are other ionic currents due to nerve depolarisation but its magnitude is too low to be detected by the skin electrode.

3.4 Generation of Ionic Current

The cells in the muscular tissue depolarise and contract when triggered by electrical stimulus. At rest the cell membrane is impermeable to ions. Let's take the potential voltage of interstitial fluid as 0 mV, the potential of intracellular fluid will be around -90 mV. When the action potential arrives at the cell, it increases the potential beyond the threshold of -70 mV. At this stage the cell membrane becomes permeable and the exchange of ions takes place, Na^+ moves inside the cell, thereby raising the potential for a short duration to +20 mV. At that moment K^+ ions start moving out of the cell, thereby decreasing the potential simultaneously. Ca^{2+} starts moving into the cell due to which the charge of the cell still remains same and a plateau is observed. Because of this a refractory period is achieved, which gives time to the cell to wait for another action potential. Before the refractory period the cell cannot be depolarized again as its potential is already positive. Then the Ca^{2+} channels are closed, so there will be no more positive ions going inside the cell while K^+ ions are still moving out. This makes the inside of the cell more negative, so the membrane potential decreases further this phase is called repolarization.

The action potential of pacemaker cells is of different shape than that of myocardial cells. Due to the slow leak of ions across the cell membrane, its action potential increases gradually with time.

3.5 Electrocardiogram

The investigation of the ECG has been extensively used for diagnosing many cardiac diseases. The ECG is a real record on the direction and magnitude of the electrical co-motion that is generated by depolarisation and repolarisation of the atria and the ventricle. One cardiac cycle in an ECG signal consists of PQRST wave. Figure 3-3 Basic ECG wave shape with morphological features shows a sample ECG signal. The majority of the clinically useful information in the ECG is originated in the intervals and amplitudes defined by its features. The improvement of precise and rapid methods for automatic ECG feature extraction is of chief importance, particularly in the long recording where the patient is implanted with some device generally called as Holter device which records the patient's heart activity for a long duration of time which can be analysed by the doctor at a later stage.

The isoelectric line is also called the baseline. It is not perfectly straight, but wandering so it is also called as wandering baseline. Which can be caused by patient movements, movement of electrode wires. The isoelectric does not have any positive or negative charge occurring. Any deflection above the isoelectric line is considered as positive, while deflection below the isoelectric line is considered as negative. When the electrical signal propagates towards the electrode, it shows up as positive deflection on the ECG. When the electrical signal travels away from the electrode it shows up as negative deflection on the ECG. The first positive deflection in the isoelectric line is called as a P wave and the first negative deflection following the P wave is Q wave. The next positive deflection after the Q wave is R Wave.

3.6 ECG Wave Description

The ECG is described by waves, interval and segments.

- Waves are labelled using the letters P, QRS, T and U.
- Segments are defined as the time duration between different waves, for example S-T segment is the duration between S and T waves. This segment denotes the entire ventricular depolarization.
- The intervals on the time durations that includes waves and segments for example Q-T interval represent the time for ventricular depolarization and repolarization.

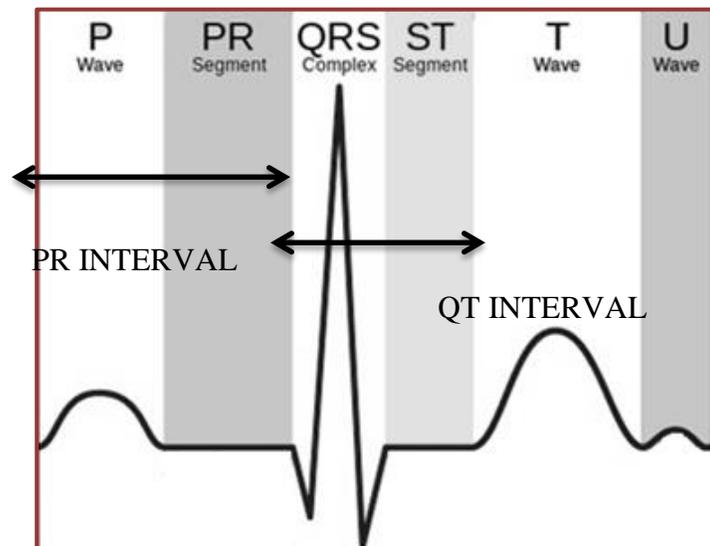


Figure 3-3 Basic ECG wave shape with morphological features

The ECG waves, interval and segments all form a feature class for the diagnosis.

- P-waves represent the depolarisation of atrial myocardium that spreads from the S-A node to the entire atria (upper heart chamber muscles); it indicates the start of atrial contractions that pumps blood to the ventricle. The duration of the P wave is usually 0.08 to 0.1 seconds.
- PR interval indicates AV conduction time, it is the time taken by the electrical impulse to travel from the sinus node through the AV node. Normally this interval is 0.12 to 0.20 seconds. The length of this interval depends on the heart rate which could be shortened during physical exercise.
- A QRS wave complex, which includes the Q, R and S waves, indicates ventricular depolarisation which results in contraction of the ventricle's that pumps blood to the lungs and other parts of the body.
- S-T segment - represents the early part of ventricular depolarization, it is the kind of the QRS wave during which the entire ventricle is depolarized. It roughly corresponds to the plateau phase of the ventricular action potential. This segment is important in the diagnosis of ventricular ischemia is under those conditions. This segment is either depressed or elevated.
- T wave - represents the repolarisation of the ventricular myocardium. It is longer in duration than depolarisation, and also slightly asymmetrical.

- Q – T interval corresponds to the time for both ventricular depolarization and repolarization. The duration of the QT interval is from the beginning of the QRS complex to the end of the T wave. Its normal range is from 0.2 to 0.4 seconds. The QT interval depends on patient gender, age and heart rate.

3.7 ECG Lead Systems

The ECG is a record of electrical activity generated by heart beats and measured from the surface of the body using special electrodes. It can be viewed, in a simpler term, as an electrical signature of heart behaviour. ECG signals are acquired by placing electrodes on the body surface at different prescribed locations and connecting the electrodes in different configurations to differential voltage amplifiers and a recorder. Three-lead ECG recording methods is based on an Einthoven Triangle named after (Malmivuo & Plonsey, 1995). Three leads are used to measure heart electrical activities, as shown in figure 3-4.

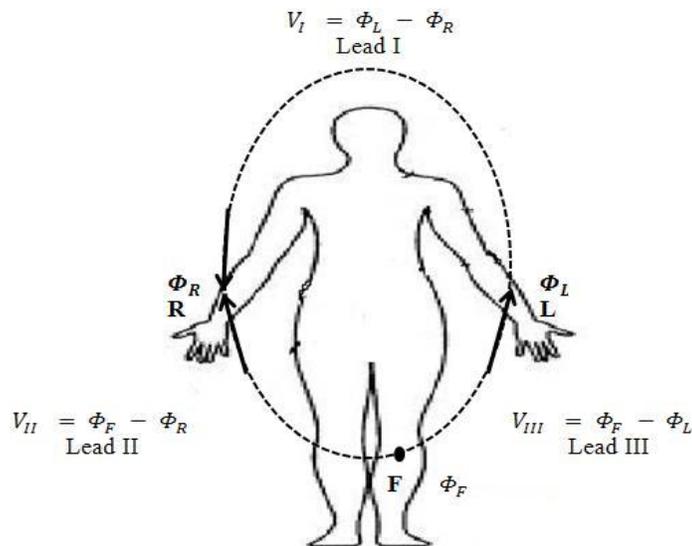


Figure 3-4 Einthoven triangle and ECG limb leads definition

The Einthoven limb leads are defined in the standard way as:

$$\begin{aligned} \text{Lead I: } V_I &= \Phi_L - \Phi_R \\ \text{Lead II: } V_{II} &= \Phi_F - \Phi_R \\ \text{Lead III: } V_{III} &= \Phi_F - \Phi_L \end{aligned} \quad (3-1)$$

Where Φ refers to measured potential from the electrodes at related locations. According to Kirchhoff's law, these lead voltages have the following relationship:

$$V_I + V_{III} = V_{II} \quad (3-2)$$

Therefore, only two of these three leads are independent.

Ever since mid-1930s, a standard 12-lead ECG system comprises of 3 limbs leads, 3 leads in which the limb potentials are referenced to a modified Wilson terminal are shown in Figure 3-5, and 6 leads placed across the front of the chest are shown in Figure 3-6 and referenced to the Wilson terminal is the standard ECG used by clinicians is the 12-lead ECG.

Limb Leads (Bipolar) – Leads I, II, III

The limb leads have electrodes attached to the limbs. Three views are acquired from Leads I, II and III, commonly refer to bipolar leads as they use only two electrodes for measurement as shown in Figure 3-1. One electrode acts as the positive electrode while the other as the negative electrode. Right Leg (RL) electrode is not used for obtaining the ECG rather than it is used as a ground reference electrode and helps to reduce instrument common mode interference. Electrically, the placement of a limb electrode in any position along the arm is the same. The minor difference is in the extra impedance of tissue resistance, if the electrode is placed further from the heart. So, an electrode can be attached either to the wrist or to the chest near that same arm.

The Augmented Limb Leads - aVL, aVR, aVF

The signals from the limb electrodes can be combined to give extra views of the heart and they are called the augmented leads. One of the limb electrodes acts as the positive electrode.

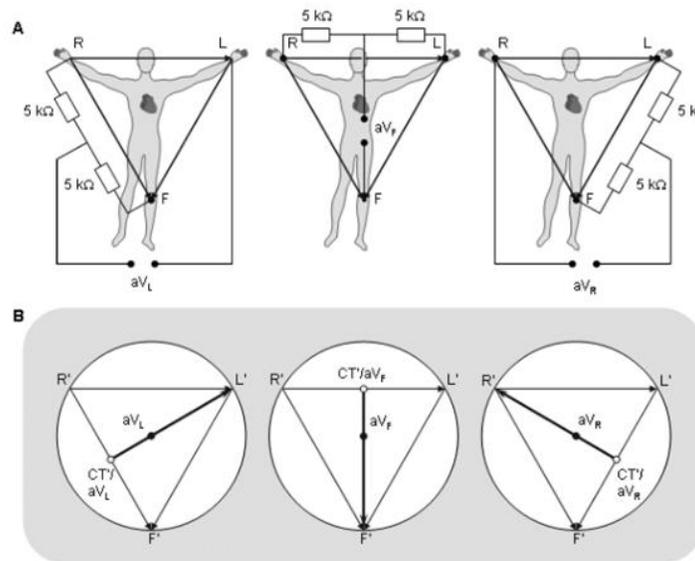


Figure 3-5 Augmented lead placement

(Figure adapted from Malmivuo & Plonsey, 1995)

The negative electrode is the average of the signals from the remaining two limb electrodes. In contrast to Leads I, II and III, the augmented leads are known as unipolar leads. In total, there are six views obtained from the limb leads, providing views from different angles along the frontal (anterior) plane.

There are six chest electrodes V_1, V_2, \dots, V_6 which gives six views of the heart signals across the front (ventral aspect) of the chest. The positive electrode is the chest electrode. The negative electrode is the Wilson Central Terminal (WCT). This is a virtual electrode with

potential obtained by averaging the signals from the three electrodes LA, RA and LL. The WCT is thus the electrical centre of the heart. These six leads are known as precordial leads and are unipolar leads.

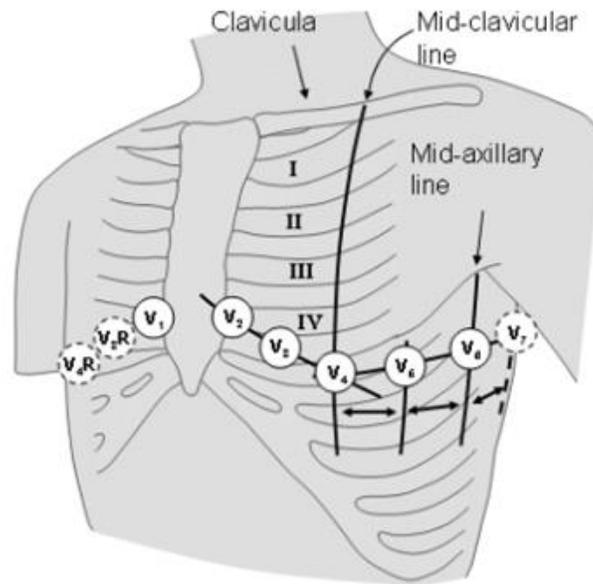


Figure 3-6 six chest electrodes placement for ECG
(Figure adapted from Malmivuo & Plonsey, 1995)

Skin or surface electrode is used to detect the depolarising signal from the excitable myocardium. When the depolarising wave propagates towards the positive electrode of the amplifier, the voltage is recorded as positive and is represented by an upward deflection in the ECG. Atrial deep polarisation is recorded as positive P wave leads I due to the propagation of an impulse from the right atrium towards the left that is towards lead I. Repolarisation wave T_a is recorded as negative P wave in the same lead. However, for most of ECGs, QRS complex mask the atria repolarisation hence T_a wave is not distinct.

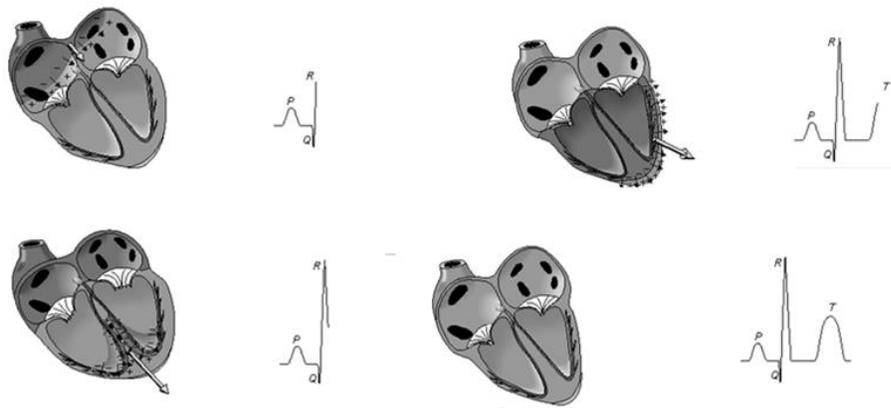


Figure 3-7 Different component of ECG during depolarization of atria and ventricle.

(Figure adapted from Malmivuo & Plonsey, 1995)

The depolarisation of ventricular myocardium starts in the Inter ventricular septum from left to right generating the Q wave of the ECG. This is followed by the near simultaneous depolarisation of left and right ventricle. Although the direction of depolarizing muscles of left and right ventricle is opposite in direction, the net direction along the horizontal axis is to the left of the heart because of the thick left ventricular wall generating large electrical potential as compared to the right ventricular wall. This depolarisation process in context with ECG is shown in figure 3-7. It has been mentioned earlier that the thick left ventricle wall generates more pressure to pump the blood throughout the body as compared to the right ventricle wall which pumps the blood to the lungs for purification. Repolarisation of ventricle starts in the opposite direction, resulting in the right T wave.

3.8 Ambulatory ECG

Ambulatory ECG monitoring is used to measure electrical activity of heart while doing normal activities or walking (ambulatory). Small electrodes are placed onto chest. Wires from the electrodes are connected to a small, lightweight recorder (Holter monitor). The recorder is attached to a belt which makes it easy to wear around waist. Ambulatory ECG record the electrical activity for 24-48 hours. It is recommended for patients with transient symptoms, e.g., palpitations, light-headedness, or syncope, which are indicative of arrhythmias. Another group of patients are those at high risk of sudden death after infarction. Ambulatory monitoring is also used in patients who are on antiarrhythmic drugs and whose reaction to the therapy needs to be assessed. The reason for using ambulatory ECG is to detect arrhythmia, which in some cases lasts for only few seconds or minutes and can be easily go undetected because of nonappearance at clinic check-up. The ambulatory ECG recording technique is also referred as Holter monitoring after its inventor Norman Holter, who introduced the first portable device to record an ECG (Holter, 1961). The Committee on Electrocardiography in The American Heart Association (AHA) had recommended the position of electrode placement for two-channel recording (Sheffield et al., 1985). They proposed a five-electrode System, one of which is a ground electrode with the other two pairs of electrode each forming a bipolar lead. The P waves are often masked by noise and artefacts, which makes diagnosis of atrial arrhythmias using the ambulatory ECG difficult by using algorithms for P wave detection (Khawaja, 2006).

3.9 Heart Rhythms and Arrhythmias

Arrhythmia is used as a collective term for all sorts of disorder in which normal sinus rhythm (NSR) is disturbed. NSR is the normal value is 60 to 100 beats per minute, with a regular R-R interval. Normal heart beat originate in the sino-atrial node, the heart's pacemaker. This is characterized with normal P wave occurrence before normal QRS wave. Any deviation from NSR is called as arrhythmias, which can be life threatening. If the heart rate is too slow, blood supply to body organ will be affected. On the other hand, if the heart rate is fast, the contraction of the ventricle occurs before getting completely filled. Thus low pumping efficiency and adverse perfusion can occur. In this study are eight arrhythmias discussed which are more frequent and all of these eight arrhythmias are used for feature extraction and classification.

3.9.1 Right Bundle Branch Block (RBBB)

Conduction system in the ventricle starts with the A-V node propagating through the Bundle of His. From there it divide into left and right bundle branches. The whole conduction system may be affected by the block in conduction from the A-V node. Myocardium depolarization is delayed due to delay in arrival depolarizing impulse. The effect of this delay is entire ventricular depolarization takes longer time, reflected by wide QRS complex and abnormal shape. In the RBBB, electrical impulse does not depolarise right ventricular myocardium due to blockage of signal in Purkinje network. Instead the electrical impulse, arrive at right ventricular myocardium from the side of left ventricular myocardium, which results in morphological change in QRS complex. The QRS complex widens more than 0.12 seconds,

in case of complete blockage. While during incomplete blockage it is around 0.10 to 0.11 seconds.

3.9.2 Left Bundle Branch Block (LBBB)

In LBBB, the electrical impulse propagation to the left ventricular myocardium is blocked, so the depolarisation of left ventricular myocardium occurs in an abnormal way. The right ventricle is depolarized first, and then the impulse travel to the left ventricular myocardium.

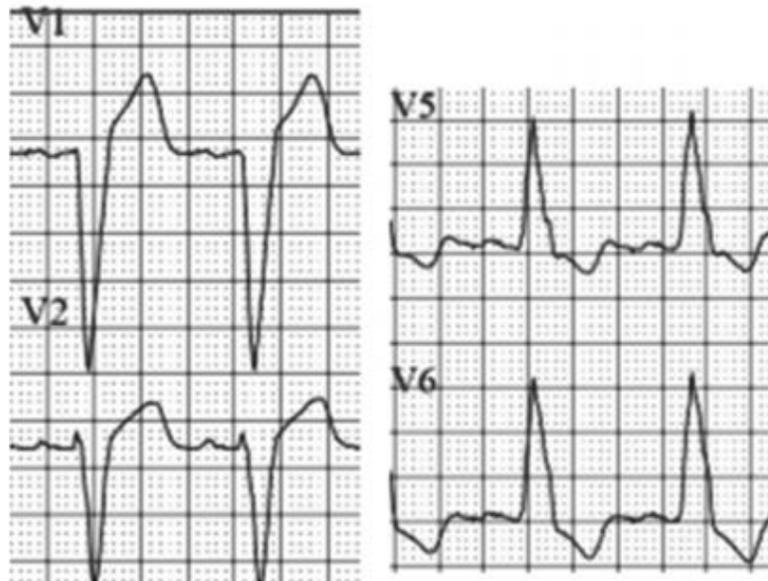


Figure 3-8 RBBB arrhythmia recorded by different leads

(Figure adapted from Rajendra Acharya, 2007)

The signal propagation through myocardium is slower visible in figure 3-8, as compared to propagation through special conducting cells.

3.9.3 Premature Ventricular Contraction (PVC)

Premature ventricular contractions (PVCs) are extra, abnormal heartbeats that begin from the ventricles. They are called as premature because of their earlier occurrence than the normal cycle. The extra pacemaker may either be in the Purkinje fibre network, Bundle of His or ventricular myocardium. The P wave maintains its rhythm and shape as seen in figure 3-9, PVCs do not depolarise the atria.

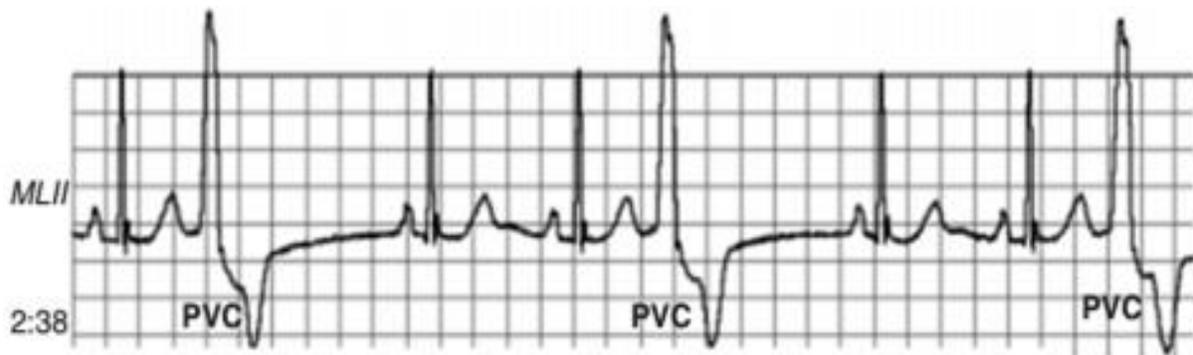


Figure 3-9 Premature Ventricular Contraction

(Figure adapted from Rajendra Acharya, 2007)

If three or more consecutive PVCs appear then this is called as Ventricular Tachycardia (VT) which is characterized by unusual wide QRS complex > 0.14 seconds. The strange shape of the QRS complex is observed with different direction from the normal QRS complex.

3.9.4 Paced Beat

This is the beat from the artificial implanted pacemaker, which generate beat around 60 to 70 per minute. It may vary according to the setting on the device. It is used by people with a

dangerously slow heartbeat. The electrode is placed on the top of the right ventricular cavity or the right atrium. The paced beat is characterized by narrow spike followed by P' wave.

3.9.5 Atrial Premature Beat

An atrial premature beat is characterised by the presence of P' wave, followed by a QRS complex and a T wave. This occurs because of an ectopic pacemaker generates impulse before the S-Anode. These impulses lead the S-Anode normal impulse and spread throughout the myocardium causing it to contract prematurely. If atrial premature beat occurs very early before the full repolarisation of bundle branches, it results in a broad QRS to bundle branch block. When three or more APBs focus simultaneously, the rhythm is considered to be atrial tachycardia. The heart rate is fast from 160 to 240 bpm.

3.9.6 Flutter Wave

Flutter wave occurs in the atria of the heart and signifies an abnormal heart rhythm with beats ranging from 240 to 360 per minute. This rhythm is common in patient having history of cardiovascular disease, e.g. hypertension, coronary artery disease and diabetes. The P' wave appear quickly and regularly, and the waveform appears like a saw tooth wave. Cardiac blood output may reduce by as much as 25% due to incomplete filling of the ventricle.

3.9.7 Ventricular Escape Beat

Ventricular escape beat occurs due to delay in the rate of electrical impulses from the S-Anode to the AV node. In VEB ectopic pacemaker in the bundle branches or Purkinje fibre network dominates at a rate of less than 40 bpm. Because of slow heart rate, the cardiac

output is also reduced, leading to hypotension and decrease blood supply to the brain and other vital organs which may result in fainting, shock or congestive heart failure.

3.10 Noise in ECG

Generally ECG signals are often contaminated with noise and artefacts of various sources . Noise is defined as presence of persistent contaminant (such as power line interference, muscle movements) while artefacts is defined as the presence of a transient interruption (such as electrode movement). (Clifford, Azuaje, & McSharry, 2006) (Rajendra Acharya, 2007) has classified various ECG noises and artefacts.

1. Power Line Interference: Power line interference is comprised of 60/50 Hz pickup and harmonics that can be modelled as sinusoids and combination of sinusoids. 50 \pm 0. 2 Hz mains noise (or 60 Hz in many datasets) with an amplitude of up to 50% of full scale deflection (FSD), the peak-to-peak ECG amplitude.

2. Electrode Contact Noise: Electrode contact noise is transient interference caused by loss of contact between the electrode and the skin, revealing sharp changes which can be permanent or intermittent. The switching action can result in large artefacts since the ECG signal is usually capacitively coupled to the system. This type of noise can be modelled as a randomly occurring rapid baseline transition that decays exponentially to the baseline and has a superimposed 60 Hz component.

3. Patient–Electrode Motion Artefacts: Movement of the electrode away from the contact area on the skin, leading to variations in the impedance between the electrode and skin, causing

potential variations in the ECG and usually manifesting themselves as rapid (but continuous) baseline jumps or complete saturation for up to 0.5 seconds;

4. Electromyographic (EMG) Noise: Electrical activity due to muscle contractions lasting around 50 Ms between DC and 10,000 Hz with an average amplitude of 10% FSD level;

5. Baseline Drift: Usually from respiration with an amplitude of around 15% peak-to-peak ECG amplitude at frequencies drifting between 0.15 and 0.3 Hz. Baseline drift cause difficulties in peak detection, as T peak could become higher than R peak and the peak detection algorithm take a T peak as R peak. Baseline drift should be removed from the ECG before feature extraction.

6. Data Collecting Device Noise: Artefacts generated by the signal processing hardware, such as signal saturation;

7. Electrosurgical Noise: Noise generated by other medical equipment present in the patient care environment at frequencies between 100 kHz and 1 MHz, lasting for approximately 1 and 10 seconds;

Conclusion

This chapter provides introduction about the heart and ECG for readers from other fields. The basic anatomy of heart is discussed along with the generation of ionic currents in the heart and their path. Various arrhythmias are discussed which were used for classification later in the thesis. Different kind of noise sources and their characteristics are also highlighted. QRS segment contains maximum information of ECG. For effective automatic ECG noise removal

and classification system feature extracted from QRS segment should be used for achieving better results.

Chapter 4 Independent Component Analysis

Independent Component Analysis (ICA) is a statistical methods that is used to identify underlying factors or component that are statistically independent (Hyvärinen & Oja, 2000a). In the last decade it has got wider attention for its application for feature extraction and blind source separation (BSS) in biomedical signal processing. The special aspect of ICA is to separate the signal mixture to its sources with no information about the original source or the mixing parameters. Though the source estimation with ICA have certain indeterminacies like

4.1 Independence

A key concept that constitutes the base of Independent Component Analysis is statistical independence. It can be understood more easily with the following example of a coin tossing. For an unbiased coin whose probability of getting tail $P_t = 0.5$, then the probability of getting two tails is

$$\begin{aligned} P_t \times P_t &= \prod_{i=1}^2 P_t && (4-1) \\ &= 0.5^2 \\ &= 0.25 \end{aligned}$$

Where the symbol \prod is standard notation for representing products, similarly the probability of obtaining exactly N tails from N coin tosses is

$$\prod_{i=1}^N P_t = P_t^N \quad (4-2)$$

The probability of obtaining a number of tails can be obtained as the product of the probability of obtaining each tail only because the outcome of the coin tosses are independent events. To simplify the above discussion consider the two different random variables x_1 and x_2 . The random variable x_1 is independent of x_2 . If the information about the value of x_1 does not provide any information about the value of x_2 and the same holds true the other way round, then it can be said that x_1 and x_2 could be random signals originating from different physical process that are not related to each other.

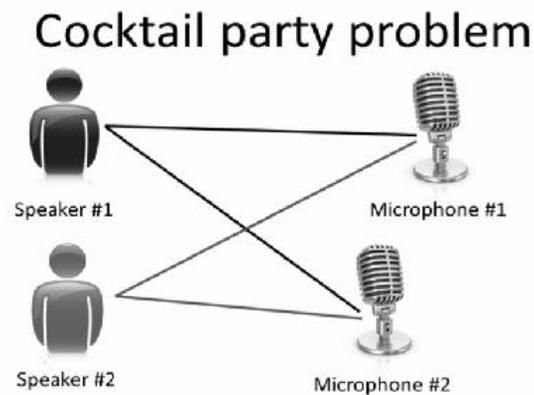


Figure 4-1 ICA demonstration with Cocktail Party Problem

The cocktail party problem is a classic example to understand blind source separation, where the objective is to separate the individual voice of the speaker from a sample of mixture of spoken voices recorded in the microphones. As shown in figure 4-1, Given the conditions of the same number of sources as there are receivers.

$$x_1(t) = a_{11} s_1(t) + a_{12} s_2(t) \quad (4-3)$$

$$x_2(t) = a_{21} s_1(t) + a_{22} s_2(t) \quad (4-4)$$

Where x_1 and x_2 are the sound signals received by the microphone. The value of signals received by each microphone depends on certain variables and in this case it is the distance of microphone from each speaker. Goal is to separate individual speaker voice from the voice mixture, with no information about the source available. Figure 4-2 demonstrate general block diagram of ICA as blind source separation technique.

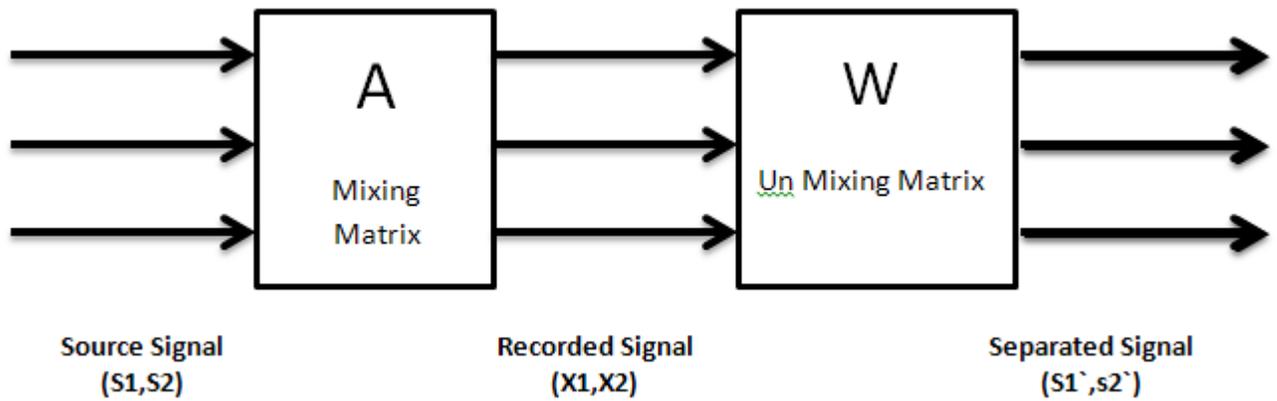


Figure 4-2 Blind source separation (BSS) block diagram

4.2 Illustrative Example of ICA

To explain the concepts discussed in the above section two simple illustrations of ICA is presented here. The results presented below are obtained using the FastICA algorithm (Hyvarinen, 1999). Other algorithms (Bell & Sejnowski, 1995) (Jean-François Cardoso & Souloumiac, 1993) could also be used but since this work is intend to use FastICA for artefact removal and classification in the later stage so FastICA is the natural choice.

In this illustration two independent signals, s_1 and s_2 , are generated. Both the signals are of the same length and zero mean. These signals are shown in Figure 4-3. The independent components are then mixed according to mixing equation (4-3)(4-4) using an randomly chosen mixing matrix A, where

$$A = \begin{pmatrix} 0.4839 & 0.7685 \\ 0.8621 & -0.5724 \end{pmatrix} \quad (4-5)$$

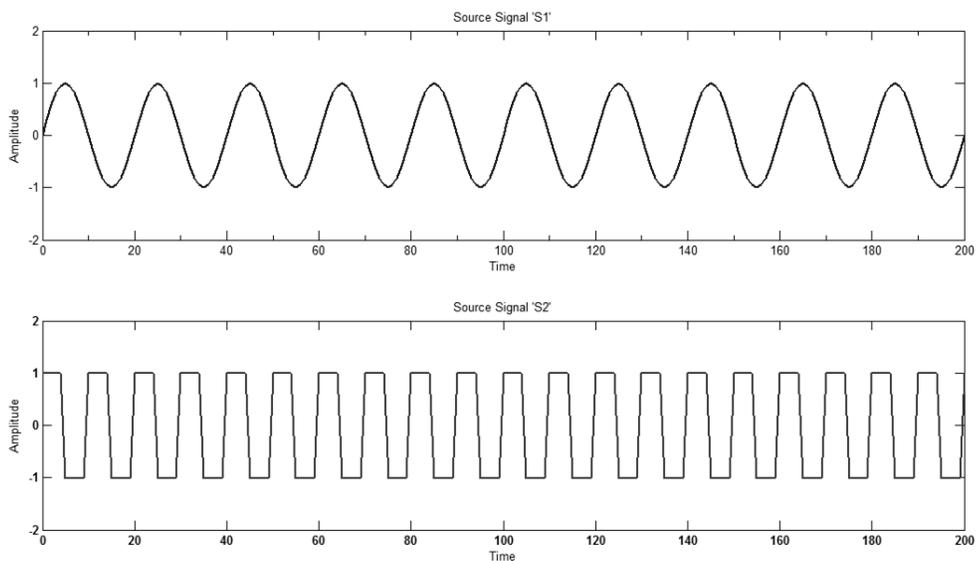


Figure 4-3 ICA demonstration original source signal

The mixing result is shown in Figure 4-4 ; it is clear from the figure that both figures are unrecognizable.

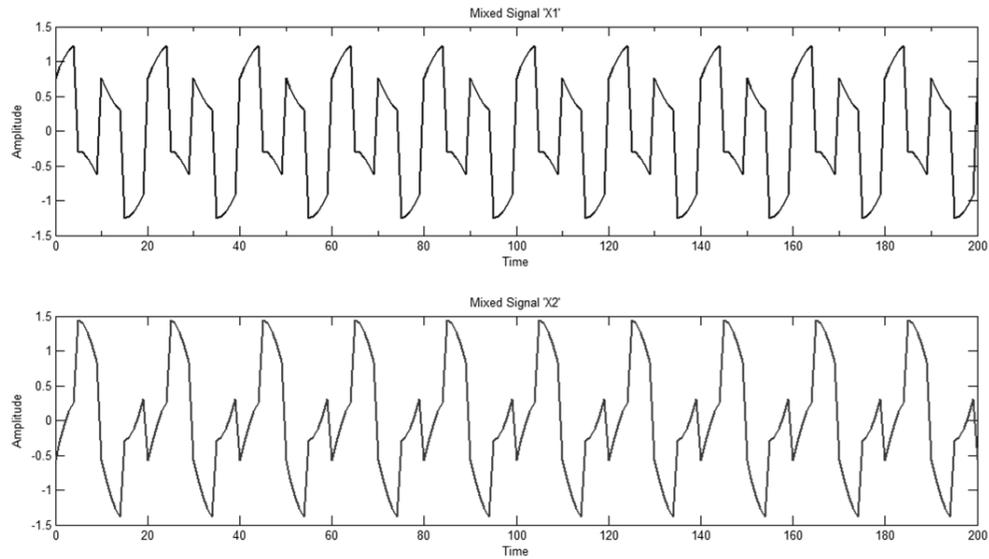


Figure 4-4 ICA demonstration mixed (observed) signal

Finally, the mixtures x_1 and x_2 are separated using ICA to obtain s_1 and s_2 , using estimated de-mixing matrix W , the separated signals are shown in Figure 4-5.

$$W = \begin{pmatrix} 0.9176 & -0.5150 \\ -0.8637 & -1.1596 \end{pmatrix} \quad (4-6)$$

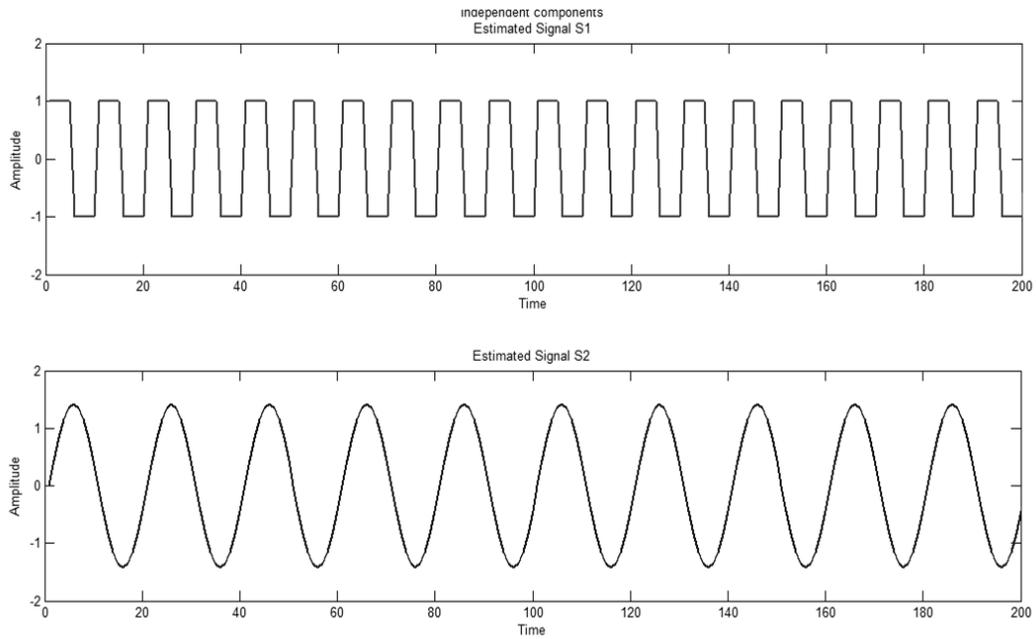


Figure 4-5 ICA demonstration recovered source signal

By comparing Figure 4-3 to Figure 4-5 it is evident that the independent components have been estimated accurately and that the independent components have been estimated without any previous knowledge of the components themselves or the mixing process.

This example also provides a clear illustration of the scaling and permutation uncertainties discussed later. The scale of the corresponding waveforms in Figure 4-3 and Figure 4-5 are different. Thus the estimates of the independent components are some multiple of the independent components of Figure 4-3. The order of the independent components has been reversed between Figure 4-3 and Figure 4-5. These uncertainty problems are rather minor when compared with the usefulness of ICA.

Ambiguities & Assumptions in ICA

There are certain inherent ambiguities in ICA model.

- 1) The variance (energies) and scale of the independent components cannot be determined. It could be explained by rewriting the mixing equation

$$x = As \quad (4-7)$$

$$= \sum_{i=0}^N a_i s_i \quad (4-8)$$

Where a_i denotes the i_{th} column of the mixing matrix A . Since both s and A being unknown any scalar multiplier in one of the source s_i could be cancelled by dividing the corresponding column a_i of A the same scalar value. But this ambiguity can be dealt with easily by considering each source has a unit variance: $E \{s_i^2\} = 1$. Then the matrix A will be adapted in ICA solution to consider this restriction. Secondly ambiguity regarding the signs of the source can be dealt easily by multiplying it by -1 without affecting the model and results.

- 2) The order of independent components cannot be determined. Formally introducing a permutation matrix P and its inverse into the mixing model to give. Here the elements of P s are the original independent component variables, but in another order.

Assumptions

- The sources being considered are statistically independent

The first assumption that is fundamental to ICA is statistical independence, which is the key feature that enables estimation of the independent components $\hat{s}(t)$ from the observations $x_i(t)$.

- The independent components have non-Gaussian distribution

The second assumption is necessary because of the close link between Gaussianity and independence. It is impossible to separate Gaussian sources using the ICA framework because the sum of two or more Gaussian random variables is itself Gaussian. That is, the sum of Gaussian sources is indistinguishable from a single Gaussian source in the ICA framework, and for this reason Gaussian sources are discouraged. This is not an overly restrictive assumption as in practice; most sources of interest are non-Gaussian.

- The mixing matrix is invertible

The third assumption is straightforward. If the mixing matrix is not invertible, then clearly the unmixing matrix estimation is not possible since it does not even exist. ((Hyvärinen & Oja, 2000a; Naik & Kumar, 2011))

4.3 Probability Density Function

Probability density function (pdf) of a continuous random variable is a function that defines the probability that the random variable takes a value in a given interval

$$p_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \left(\frac{(x - \bar{x})^2}{2\sigma^2} \right) \quad (4-9)$$

4.4 Central Limit Theorem

The central limit theorem states that the mean of a sufficiently large number of iterates of independent random variables, each with a well-defined mean and well-defined variance, will be approximately normally distributed under certain conditions (Rice, 2007).

If a set of signals $s = (s_1, s_2 \dots s_N)$ are independent with means $(\mu_1, \mu_2 \dots \mu_N)$ and variances $(\sigma_1^2, \sigma_2^2, \dots, \sigma_N^2)$ then, for a large number N of signals s , the signal

$$x = \sum_{i=1}^N s_i \quad (4-10)$$

Has a probability distribution function which is approximately Gaussian. The Central Limit Theorem does not place restrictions on how much of each source signal contributes to a signal mixture, so above description holds good even if the mixing coefficients are not equal to unity.

$$x = \sum_{i=1}^N s_i a_i \quad (4-11)$$

Where a_i 's are non-zero mixing coefficients. We can say that from a mixture formed using mixing coefficients $a = a_1$ and $b = b_1$

$$x = as_1 + bs_2$$

4.5 Moments

Moments are measures that are used to study the distribution and it provides a great deal of information in numerical form about the distribution.

First Moment

The first moment of pdf gives mean value \bar{x} of the signal x . It is a measure of central tendency sometime it is also called as location.

$$E[x] = \int_{x=-\infty}^{+\infty} p_x(x)x dx \quad (4-12)$$

Second Moment

$$E[x^2] = \int_{x=-\infty}^{+\infty} p_x(x)x^2 dx \quad (4-13)$$

Third Moment

$$E[x^3] = \int_{x=-\infty}^{+\infty} p_x(x)x^3 dx \quad (4-14)$$

$$E[(x - \bar{x})^3] = \int_{x=-\infty}^{+\infty} p_x(x)(x - \bar{x})^3 dx \quad (4-15)$$

Fourth Moment

The *fourth moment* $E[x^4]$ if the probability density function is

$$E[x^4] = \int_{x=-\infty}^{+\infty} p_x(x)x^4 dx \quad (4-16)$$

The different moments discussed before needs an infinite number of signal values. For a finite number of N sampled points, these can be defined as

The first moment $E[x]$ of a signal x is calculated as the mean

$$E[x] = \bar{x} \quad (4-17)$$

$$\approx \frac{1}{N} \sum_{t=1}^N x^t \quad (4-18)$$

$$E[x^2] = \bar{x}^2 + \sigma^2 \quad (4-19)$$

If the variable x has zero mean that is ($\bar{x} = 0$) then the second moment is equal to the variance

$$E[x^2] = \sigma^2 \quad (4-20)$$

The probability distribution function of \mathbf{x} is approximately Gaussian. This implies that independent random variable is more non Gaussian than their mixture. Hence, non gaussianity is a measure of independence and forms one of the bases for independent component analysis.

Non-Gaussianity is an important and essential principle in ICA estimation. To use non Gaussianity in ICA estimation, there needs to be a quantitative measure of non Gaussianity of a signal. Before using any measures of non Gaussianity, the signals should be normalized. Some of the commonly used measures are kurtosis and entropy measures.

Kurtosis is the classical method of measuring Non Gaussianity (Naik & Kumar, 2011; Tanskanen, Mikkonen, & Penttonen, 2005). When data is pre-processed to have unit variance, kurtosis is equal to the fourth moment of the data. The Kurtosis of signal (s), denoted by Kurt (s), is defined by equation (4-21)

$$kurt(s) = E\{s^4\} - 3(E\{s^2\})^2 \quad (4-21)$$

This is a basic definition of kurtosis using higher order (fourth order) cumulant; this simplification is based on the assumption that the signal has zero mean. To simplify things, it can be further assumed that (s) has been normalized so that its variance is equal to one: $E\{s^2\} = 1$. Kurtosis has been widely used as measure of Non Gaussianity in ICA and its related fields because of its computational and theoretical simplicity (Naik and Kumar, 2011).

Entropy is a measure of the uniformity of the distribution of a bounded set of values, such that a complete uniformity corresponds to maximum entropy. From the information theory

concept, entropy is considered as the measure of randomness of a signal. Entropy H of discrete-valued signal S is defined as

$$H(S) = -\sum P(S = a^i) \log P(S = a^i) \quad (4-22)$$

One fundamental result of information theory is that Gaussian signal has the largest entropy among the other signal distributions of unit variance. Entropy will be small for signals that have distribution concerned with certain values or have a pdf that is very "spiky". Hence, entropy can be used as a measure of non-Gaussianity. In ICA estimation, it is often desired to have a measure of non-Gaussianity which is zero for Gaussian signal and nonzero for non-Gaussian signal for computational simplicity. Entropy is closely related to the code length of the random vector. A normalized version of entropy is given by a new measure called Negentropy J which is defined as

$$J(S) = H(s_{gauss}) - H(s)$$

Where s_{gauss} is the Gaussian signal of the same covariance matrix as (s) . Above equation shows that Negentropy is always positive and is zero only if the signal is a pure Gaussian signal. It is stable, but some times it is difficult to calculate it. Hence approximation must be used to estimate entropy values.

4.6 Preprocessing

To simplify the ICA algorithms, the following Preprocessing steps, namely, Centering and Whitening/Sphering are used:

4.6.1 Centering

Centering is a simple preprocessing operation that is commonly performed is to “centre” the observation vector x by subtracting its mean vector $E\{x\}$. New vector, x_c , is obtained with equation (4-23), where \bar{x} is the mean of x .

$$x_c = x - \bar{x} \quad (4-23)$$

In Geometrical terms, subtracting the mean is equivalent to moving the centre of coordinates to the origin. The mean can be re-added to the final result at the end. After the unmixing matrix has been estimated using the entered data, the next step is to obtain the actual estimates of the independent components using the equation (4-24)

$$\hat{s}(t) = A^{-1}(x_c + \bar{x}) \quad (4-24)$$

4.6.2 Whitening

Whitening operation removes all linear dependencies in a data. It involves linearly transforming the observation vector such that its components are uncorrelated and have unit variance (Karhunen & Oja, 2001; Meyer, 2000; Stone, 2004). Let x_w Indicate the whitened vector, then it satisfies the following equation:

$$E\{x_w x_w^T\} = I \quad (4-25)$$

where the covariance matrix of x_w is given by $\{x_w x_w^T\}$, whitening is also termed as sphering, where whitening maps the data into a spherically symmetric distribution. (Meyer, 2000) proposed a simple method to perform the whitening by using the eigenvalue decomposition (EVD) of matrix x . By using EVD covariance matrix x can be modified as

$$E\{x x^T\} = V D V^T \quad (4-26)$$

where V is the eigenvector matrix of $\{x x^T\}$, and D is the diagonal matrix of eigenvalues. $D = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, whitened vector is given by equation (4-27)

$$x_w = V D^{-1/2} V^T x \quad (4-27)$$

Whitening transforms the mixing matrix into a new one, which is orthogonal

$$x_w = V D^{-1/2} V^T A s \quad (4-28)$$

Whitening can be expressed as

$$z = x_w x = x_w A s \quad (4-29)$$

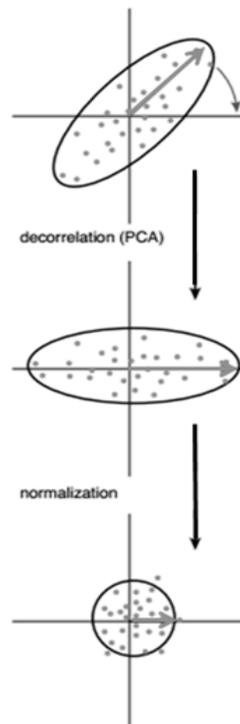


Figure 4-6 Whitening Process

Adapted and modified from (Shlens, 2014)

The process of whitening a data set can be explained as a series of two linear operations. Data is projected on the principal components, as shown in figure 4-6. Each axis is then scaled so that every direction has unit variance; the arrow indicates the transformation of the eigenvector with largest variance.

4.7 ICA Algorithms

This section provides description of various ICA algorithms used in this study. Key points related to ICA application in noise removal from ECG are discussed here for more details reader can go to the references discussed in this section.

4.7.1 Infomax ICA

The methods used to find the de-mixing differentiate the ICA algorithms. Infomax ICA and FastICA are the two major types of ICAs according to their distinctive mechanisms of operation.

Infomax algorithm attempts to separate signals through minimizing mutual information or maximizing entropy in outputs signals via unsupervised learning. The algorithm is often deemed as a single layered neural network. The ICA neural networks are unsupervised models for separating statistically independent signals and are previously applied to speech separation problems by Bell and Sejnowski (Bell & Sejnowski, 1995).

The mutual information measures how far the two variables are from statistical independence. The multi information, a generalization of mutual information, measures the statistical dependence between multiple variables (Shlens, 2014).

$$I(y) = \int P(y) \log_2 \frac{P(y)}{\pi_i P(y_i)} dy \quad (4-30)$$

It is a non-negative quantity, which has a minimum value of zero only in the case when two variables are statistically independent. The objective now is to find a rotation matrix V such that $I(\hat{s}) = 0$. V is a rotation matrix and in two dimensions V has the form

$$V = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad (4-31)$$

Only free variable is the rotation angle θ

The learning objective of Infomax ICA algorithm is to minimize the mutual information between the outputs as illustrated in Figure 4-7.

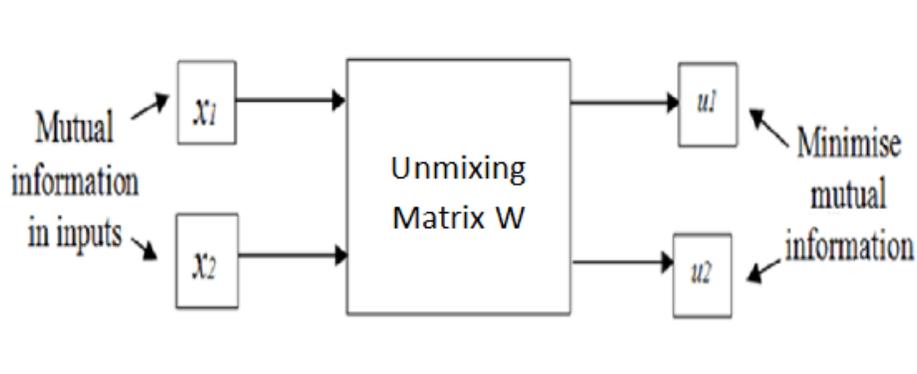


Figure 4-7 Learning objectives of Infomax ICA algorithm

Two independent sources, s_1 and s_2 , are linearly mixed by arbitrary coefficients a_{11} , a_{12} , a_{21} , and a_{22} to give the mixture x_1 and x_2 according to Equations 4-3 and 4-4. When written in a

matrix format $x = As$, where the input vector is $S = [s_1, s_2]^T$, the mixture vector is

$X = [x_1, x_2]^T$ and the mixing matrix A is

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (4-32)$$

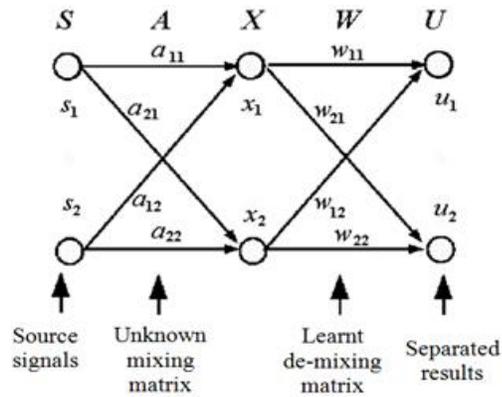


Figure 4-8 Mixing and de-mixing model

The ICA to find a de-mixing matrix W as shown in the figure 4-8, so that u_1 and u_2 , which are recovered versions of s_1 and s_2 , can be obtained by

$$U = WX \quad (4-33)$$

Where the de-mixing matrix is

$$W = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \quad (4-34)$$

And the recovered vector is

$$U = [u_1, u_2]^T \quad (4-35)$$

This is achieved by minimising mutual information found in \mathbf{u}_1 and \mathbf{u}_2 by using an unsupervised neural network with only one linear summation layer as depicted in Figure 4-8. The two neurons have linear summation basis functions, but may have different types of activation functions. The following activation function proposed by (Madhuranath & Haykin, 1998), is used as a starting point. The activation might be optimized, as per the requirement and condition of the signal.

$$f(z) = \frac{1}{2}z^5 + \frac{2}{7}z^7 + \frac{15}{2}z^9 + \frac{2}{15}z^{11} - \frac{112}{3}z^{13} + 128z^{15} - \frac{512}{3}z^{17} \quad (4-36)$$

Where z is used to represent the summed input signals being sent to the activation function.

The training follows weight updating formula .

$$W(n+1) = W(n) + \alpha [I - f(U(n))U^T(n)]W^{-T}(n) \quad (4-37)$$

Where α is the step size.

4.7.2 FastICA Algorithm

FastICA developed by Hyvärinen, 1999. (Hyvärinen, 1999) is another possible algorithm for independent component analysis. It uses maximum non-Gaussianity as a criterion of statistical independence. The algorithm is based on the central limit theorem which is used as a measure of independence. FastICA is a fixed point ICA algorithm that employs higher order statistics for the recovery of independent sources. FastICA can estimate ICs one by one or simultaneously (symmetric approach). FastICA uses simple estimates of Negentropy based on the maximum entropy principle to measure non gaussianity. This can be defined as:

$$J(\mathbf{x}) = H_G(\mathbf{x}) - H(\mathbf{x}) \quad (4-38)$$

Where \mathbf{X} is a random vector known to be non-Gaussian, $H(\mathbf{x})$ is the entropy and $H_G(\mathbf{x})$ is the entropy of a Gaussian random vector whose covariance matrix is equal to that of \mathbf{x} . For a given covariance matrix, the distribution that has the highest entropy is the Gaussian distribution. Negentropy is thus a strictly positive measure of non-Gaussianity. (Hyvärinen & Oja, 2000b) proposes some modification of the above methods for calculation of negentropy

$$J(V) = E(\phi(V)) - E(\phi(U))^2 \quad (4-39)$$

Where V is a standardized non-Gaussian random variable (zero mean and unit variance), U a standardized Gaussian random variable and $\phi(\cdot)$ a non-quadratic function (generally Tanh (\cdot)).

After some modifications FastICA algorithm can be explained in these steps.

1. Let $i = 0$ initialize the weight vector: $w = w(0)$
2. Increment i ; $i = i + 1$
3. Adjust w $w(i + 1) = E\{zg(w_i^T z)\} - E\{zg'(w_i^T z)\} w_i$
4. Normalize $w(i + 1) = \frac{w(i+1)}{\|w(i+1)\|}$
5. If convergence is not achieved return to step 3
6. After getting convergence find independent component $y_1 = wZ$

Where $Z = z_1, z_2, z_3 \dots z_n$ is whitened signal matrix and $Y = y_1, y_2, y_3 \dots y_n$ are estimated independent components.

4.7.3 JADE Algorithm

The JADE algorithm is originally developed and implemented by Jean-Francois Cardoso and Antoine Souloumiac (Jean-François Cardoso & Souloumiac, 1993; Jean-Francois Cardoso & Souloumiac, 1996). JADE has been successfully applied to the processing of real data sets, such as found in mobile telephony and in airport radar as well as in bio-medical signals (ECG, EEG, and multi-electrode neural recordings) (Jean-Francois Cardoso & Souloumiac, 1996; Cichocki, Amari, Siwek, Tanaka, & Phan, 2003).

The JADE algorithm uses the second and fourth order cumulant in order to separate the source signals from mixed signals. The second order cumulant is used to ensure the data is “white” (i.e. decorrelated). This produces a whitening matrix \widehat{W} and the whitened sources. A set of cumulant matrices is estimated from the whitened sources. The JADE contrast function is the sum of squared fourth order cross cumulant.

- Estimation of whitening matrix.
- Estimate a maximal set $\{\widehat{Q}\}_i^Z$ of cumulants matrices.
- Optimize an orthogonal contrast. Find the rotation matrix \widehat{V} Such that the cumulant matrices are as diagonal as possible, that is, achieved by solving
$$\widehat{V} = \arg \min \sum_i \text{Off}(V^T \widehat{Q}_i^Z V).$$
- Estimate A as $\widehat{A} = \widehat{V}W^{-1}$ or estimate the components as $\widehat{S} = \widehat{A}^{-1}X$.

(Matei, 2000) concludes FastICA algorithm has better performance as compared to JADE. (Li, Powers, & Peach, 2000) also demonstrated that the fixed-point algorithms are the fastest than the JADE algorithm, and tended to produce better SNR.

Conclusion

In this chapter the basic principle behind the independent component analysis technique is discussed. The contrast functions for different routes to independence are clearly depicted. Different existing algorithms for ICA are briefly illustrated and are critically examined with special reference to their algorithmic properties. The ambiguities present in these algorithms are also presented. Comparison of different ICA algorithms is made with explanation for the selection of FastICA algorithm for noise removal and feature extraction in this study. Some of the futuristic works on ICA technique which need further investigation are development of nonlinear ICA algorithms and improvement of permutation and scaling ambiguities existing in present ICA.

Chapter 5 Principal Component Analysis

Principal Component Analysis (PCA) is a method which is commonly used in multivariate statistical analysis. Its objective is to reduce the number of dimensions from a multi variable measurement. Due to this dimensional reduction, this PCA looks for simplifying a statistical problem with the minimal loss of information. This method is also used in signal processing for separating a linear combination of signals generated from sources that are statistically independent (Castells, Laguna, Sö, Bollmann, & Roig, 2007; Inaki Romero, 2010). Where data are represented in a new coordinates having maximum variance with each other.

Many researchers have compared the capabilities if ICA and PCA for removing artefacts from various signals EEG , ECG etc. (Jung et al., 1998) in their pioneer work in removing electroencephalographic artefacts with PCA and ICA concludes that ICA appears to be a generally applicable and effective method for removing a wide variety of artefacts from EEG records , giving the preference to ICA because of it is generally applicable to removal of a wide variety of EEG artefacts . In nearly every case, ICA preserves and recovers more brain activity than PCA in both simulated and real EEG data.

For ECG noise and artefact removal (Chawla, 2009) compared PCA and ICA processing methods for removal of artefacts and noise in electrocardiograms. Chawla observed that PCA removes, the noise and artefacts contained in the ECG signals to a great extent by decorrelation and dimension reduction process but noise and artefacts of the original and the corrected ECG signals are better understood using combined PCA de-noising and ICA cleaning methods. The results conclude that higher order statistical tool like ICA and its

various versions can be effectively used for morphological feature extraction of ECG better than PCA technique by the clinicians and researchers.

PCA is a linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on (Smith, 2002). Let \mathbf{X} be the original data set, where each column is a single sample. In the \mathbf{X} is an $m \times n$ matrix where m = number of observations and n = all measurements from one particular observation. Let \mathbf{Y} be another $m \times n$ matrix related by a linear transformation \mathbf{P} . \mathbf{X} is the original recorded data set and \mathbf{Y} is a new representation of that data set.

$$\mathbf{P}\mathbf{X} = \mathbf{Y} \quad (5-1)$$

Where \mathbf{P} is a matrix that transforms \mathbf{X} into \mathbf{Y}

General method adopted to calculate PCA is

- Data Collection
- Mean Subtraction to get zero mean
- Calculation of covariance matrix
- Calculating the eigenvectors and eigenvalues of the covariance matrix
- Selecting components and forming a feature vector (such that the eigenvector with the highest eigenvalue is the principle component of the dataset)

$$\mathbf{P}\mathbf{X} = \begin{bmatrix} p_1 \\ \vdots \\ p_m \end{bmatrix} [x_1 \quad \cdot \quad x_n] \quad (5-2)$$

$$Y = \begin{matrix} p_1 \cdot x_1 & \dots & p_1 \cdot x_n \\ \vdots & \ddots & \vdots \\ p_m \cdot x_1 & \dots & p_m \cdot x_n \end{matrix} \quad (5-3)$$

It can be noted that each element of each column of Y is a dot product with the corresponding row in P , thus, the rows of P form a new set of basis vectors for representing of columns of X .

PCA utilizes the first and the second moments of the measured data, hence relying heavily on Gaussian features. ICA exploits inherently non-Gaussian features of the data and employs higher moments.

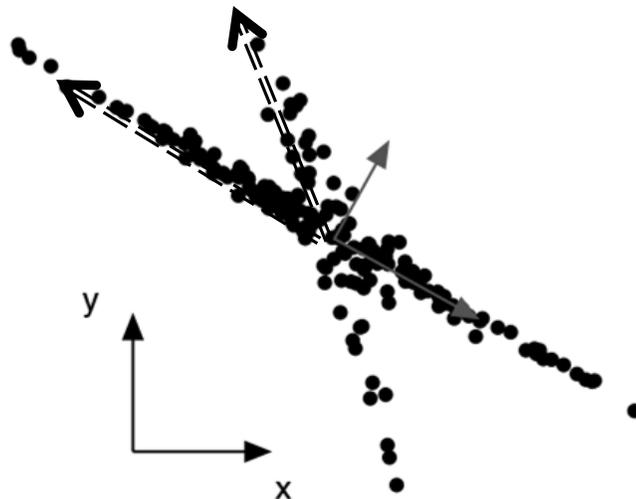


Figure 5-1 Demonstration of difference between PCA and ICA

To explain more clearly the difference between PCA and ICA consider the data distribution in Figure 5-1. The red coloured axis are principal components and blue coloured are independent components. The axes with the largest variance do not correspond to the best solution but PCA condition for finding orthogonal axis causes it to fail. Whereas there is no such

condition with independent components to be orthogonal, but to be with maximum variance for which ICA is best suited.

(Asl et al., 2008a) used PCA along with multi-layer perceptron for ECG classification into six classes of arrhythmias with training accuracy of 96.93%. More elaborate study is conducted by (Martis et al., 2013) for ECG beat classification using PCA ,LDA and ICA with the application of PCA 99.00% accuracy is achieved. Using PNN classifier of spread parameter as 0.40. Also the LDA has achieved 98.59% of accuracy with NN classifier. However the ICA yielded the highest performance with 99.28% accuracy, using PNN classifier with a spread parameter of 0.03.

In this study, principal components are selected to form a feature set along with dynamic features for arrhythmia classification. Firstly the ECG signal from MIT BIH arrhythmia database is subjected to QRS complex detection using Pan-Tompkins method. After QRS complex detection, 200 data point sample with 100 data points from the right of QRS peak, 99 data points to the left of QRS complex and the QRS peak itself is chosen as a segment of ECG beat. The choice of 200 data points around the R peak as a signal window length is such that it consists almost one cycle of cardiac activity. This duration is used in author's previous studies as well (Sarfraz, Li, & Khan, 2014; Sarfraz & Li, 2013). 100 samples of various arrhythmias are selected for extraction of principal components. The detailed procedure of PCA is explained below

- Computing the covariance matrix of the data

$$C = (X - \bar{x})(X - \bar{x})^T \quad (5-4)$$

Where X is the data matrix of 100 X 200 dimensions, \bar{x} represents mean vector of X

- Computing the eigenvector matrix V an eigenvalues diagonal matrix D with an equation (5-5)

$$D = V^{-1}CV \quad (5-5)$$

- The eigenvectors in V are arranged in the descending order of eigenvalues in D and the data is projected on these eigenvector directions by taking the dot product between the data matrix and sorted eigenvector matrix with an equation (5-6)

$$Data\ Projection = [V^T(X - \bar{x})^T]^T \quad (5-6)$$

Where V is of 200×200 dimensions, each row of it is an eigenvector. The first fifteen columns of the projected data are considered as the fifteen features for later classification.

Conclusion

In this chapter, the basic principle behind the principal component analysis technique was discussed. Important work in ECG noise removal and classification using principal component analysis were discussed. The difference between the principal component analysis and independent component analysis was explained. The general method adopted for estimation of principal components is mentioned. The limitation of principal component analysis in utilising only first and second moments and their effect in selection of important features is also discussed. The higher order statistics used in ICA make it more robust than PCA for data reduction and finding direction of maximum variance vector.

Chapter 6 Artificial Neural Network

An Artificial Neural Network (ANN) is an information processing model adapted on the line of biological nervous system. The basic unit of ANN is neuron, which form a highly interconnected grid to solve problems. ANN have been successfully used to perform various roles in various fields of application including vision, speech control, classification, identification, pattern recognition, control system and robotics.

Some characteristic of artificial neural network are

- Adaptive learning:
- Self-organization:
- Real time operation

A neuron is an electrically excitable cell that processes and transmits information through electrical and chemical signals. It consists of three parts cell body (soma), axon and dendrites. The signals are collected from another neuron through dendrites, which are thin structures of the cell body. The neuron sends out spikes of electrical activity through a long, thin cellular extension known as axon which splits into branches. At the end of each branch, a structure called synapse permits a neuron (or nerve cell) to pass an electrical or chemical signal to another cell (neural or otherwise). This electrical signal can excite or inhibit the activity in the connected neuron. If the excitatory input received is larger than in inhibitory input, the neuron responds to it by issuing a new pulse which travels along its axon on the other hand if the excitatory input is less than the threshold the neuron remain inactive. Learning occurs by changing the effective synaptic strength between neurons so that the influence of one neuron

on other changes. These features are simulated in artificial neural network model. Figure 6-1 shows general layout of ANN. The most common type of artificial neural network consists of three groups or layers of units: a layer of “input” units which is connected to a layer of “hidden” units, which is connected to a layer of “output” units. The activity of the input units represents the raw information that is fed into the network. The action of each hidden unit is defined by the activities of the input unit and the weights on the links between the input and the hidden units. The behaviour of the output units depends on the activity of the hidden units and the weights between the hidden and output units.

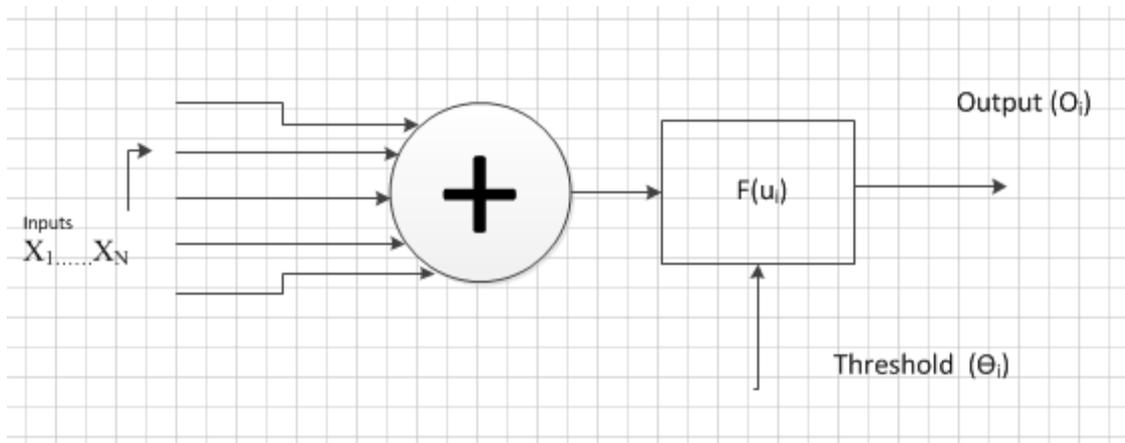


Figure 6-1 Layout of ANN

6.1 Network Performance Analysis

The performance of neural network in classification is estimated by following indices.

- Classification accuracy
- Sensitivity

- Specificity
- Positive predictive accuracy

6.2 Back Propagation Network

There are two types of learning algorithms: supervised and unsupervised. In the supervised, learning the system weights is randomly allotted at the start; and progressively modified in the light of desired outputs for a set of training inputs. The difference between the desired output and the actual output is calculated for every input, and the weights are modified in proportion to the error factor. The procedure is carried on until the system error is brought down to an acceptable limit.

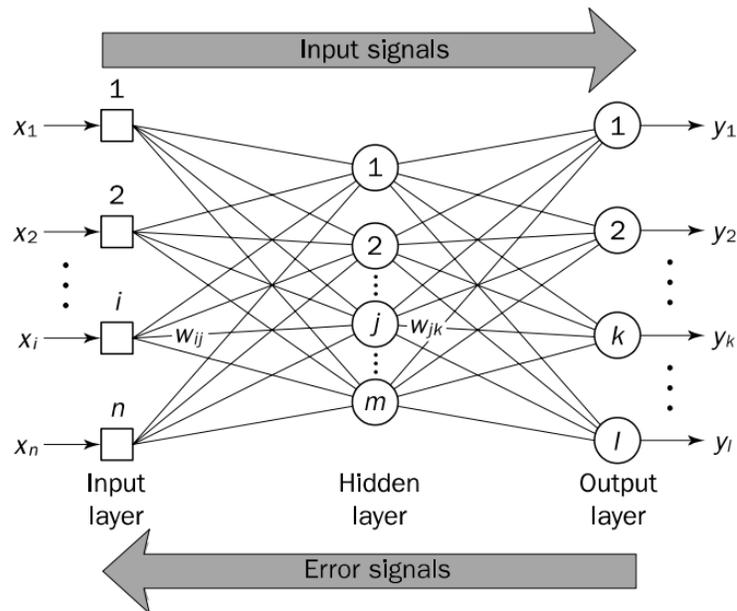


Figure 6-2 Three layer back-propagation neural network

In the figure 6-2 the indices i , j and k here refer to neurons in the input, hidden and output layers, respectively. Input signals x_1, x_2, \dots, x_n , are propagated through the network from left

to right, and error signals e_1, e_2, \dots, e_3 , from right to left. The symbol W_{ij} indicates the weight of the connection between neuron i in the input layer and neuron j in the hidden layer, and the symbol W_{jk} indicates the weight between neuron j in the hidden layer and neuron k in the output layer.

In the back propagation algorithm, objective is to reduce the overall system error to a minimum. The weight increment is directed towards the minimum system error and therefore termed as ‘gradient descent’ algorithm. There is no single rule to select the step size for the weight increment; but the step length certainly has an effect on the speed of convergence. It has been observed that for good speed, the step size should neither be ‘too large’, nor ‘too small’. In the present case, a near optimum learning constant $\eta = 0.9$ (which controls the step size), is chosen by trial and error. Since the weight increment is accomplished in small steps, the algorithm also bears the name ‘Delta Rule’

The entire operation of updating the weight matrix is a slow due to small incremental step movement towards a global minimum of a system error function. Sometimes the system gets stuck in local minima and unable to come out of it. To avoid such a scenario, the algorithm incorporates a ‘momentum term’ into its update increment (Qiu, Varley, & Terrell, 1992; C.-C. Yu & Liu, 2002). The term is a fraction of an increment of its previous step; this term tends to push the present, increment in the same direction as that of the previous step. This term helps to get out of ‘small’ dips in the path.

$$\Delta w_{jk}(p) = \beta * \Delta w_{jk}(p - 1) + \alpha * y_j(p) * \delta_k(p) \quad (6-1)$$

The equation (6-1) is a generalised delta rule (Rumelhart, Hinton, & Williams, 1986) where $\Delta w_{jk}(\mathbf{p})$ is the weight correction at p^{th} iteration, $\delta_k(\mathbf{p})$ is the error gradient at network k in the output layer , β is a momentum constant ($0 \leq \beta < 1$) , generally the momentum constant is set to 0.95.

The outputs of the hidden layer (S_j^h) and output layer (b_k) are given by the equations

$$S_j^h = f \left(\sum_{i=1}^n w_{ji}^h S_i - \theta_j^h \right) \quad (6-2)$$

$$b_k = f \left(\sum_{j=1}^n w_{kj}^o S_j^h - \theta_k^o \right) \quad (6-3)$$

Where w_{ji}^h and w_{kj}^o represent the weights corresponding to the hidden and output layers respectively and θ_j^h and θ_k^o are the biased terms of hidden and output layers respectively.

While the k^{th} component of the output error vector (\mathbf{e}_k) and hidden layer error (\mathbf{e}_j) vector:

The update equations of the output and hidden layers are given as

$$w_{kj} \text{ (new)} = w_{kj} + \eta S_j^h e_k \quad (6-4)$$

$$w_{ji} \text{ (new)} = w_{ji} + \eta S_i e_j \quad (6-5)$$

The updating of the hidden layer is more computationally intensive than the output layer. If there are more hidden layers, the computation to progressive increases. In most practical cases

a single hidden layer network is adequate (Rajendra Acharya, 2007). A change in the number of neurons in the hidden layer (n) can affect the sophistication of classification. At the same time, an increase in the number of hidden neurons may cause a delay in the convergence of weights.

Self Organising Maps

Self-organising neural networks are effective in dealing with unexpected and changing conditions (Negnevitsky, 2005). Teuvo Kohonen introduced a special class of artificial neural networks called self-organising feature maps (T Kohonen, 1987; Teuvo Kohonen, 1988). These maps are based on competitive learning. In competitive learning, neurons compete among themselves to be activated. In competitive learning only a single output neuron is active at any time. The output neuron that wins the ‘competition’ is called the winner-takes-all neuron.

Self-Organizing Maps (SOM) are applied for classification of ECG signals as these networks learn to detect regularities and correlation in their input and adapt their future responses to that input accordingly. The neurons of competitive networks learn to recognize groups of similar input vectors. Self-organizing maps learn to recognize groups of similar input vectors in such a way that neurons physically close together in the neuron layer respond to similar input vectors (T Kohonen, 1987). Each output neuron by means of these lateral connections is affected by the activity of its neighbours. The activation of the output units according to Kohonen’s original work is by equation 1. The modification of the weights is given by the equation (6-6) where O_j = activation of output unit, x_i = activation value from input unit, w_{ji}

Lateral weights connecting to output unit, \mathbf{d}_j = neurons in the neighbourhood, F_{\min} = unity function returning 1 or 0 (McGarry, Sarfraz, & MacIntyre, 2007) .

$$O_j = F_{\min}(d_j) = F_{\min} \left(\sum_i (x_i - w_{ji})^2 \right) \quad (6-6)$$

6.3 Support Vector Machine

Support Vector Machine (SVM) is an extensively used tool for binary classification problem (Vapnik, Golowich, & Smola, 1997). It is characterized by a good generalization performance. The construction of SVM classifier is based on finding a maximum margin between the training data and the decision boundary, as shown in figure 6-3. The subsets of training data which are closest to the decision boundary are called support vectors. For a given training data (x_i, y_i) for $i = 1 \dots N$, the optimization problem for the SVM is formulated as shown in equation (6-7)

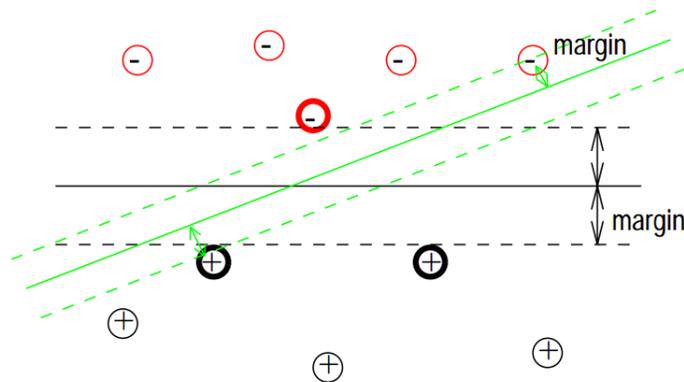


Figure 6-3 SVM methodology of maximising margin between different classes

$$A = \min_{w, \xi, b} J(w, \xi) = \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i \quad (6-7)$$

Subject to the constraints

$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, i = 1, \dots, N \quad (6-8)$$

$$\xi_i \geq 0, i = 1, \dots, N$$

Where C is a positive regularization parameter which is chosen empirically, w is the weight vector of training parameters, ξ_i is a positive slack variable indicates the distance of x_i with respect to the decision boundary, and ϕ is a nonlinear mapping function used to map input data point x_i into a higher dimensional space. SVMs can be written using Lagrange multipliers with $\alpha \geq 0$ for equation (6-8). The solution for the Lagrange multipliers are obtained by solving a quadratic programming problem. The SVM decision function can be expressed as

$$g(x) = \sum_{x_i \in SV} \alpha_i y_i K(x, x_i) + b \quad (6-9)$$

Where $K(x, x_i)$ is the kernel function and defined as

$$K(x, x_i) = \phi(x)^T \phi(x_i) \quad (6-10)$$

For mapping the data into high dimensional space, various kernel transformations namely: linear, quadratic, polynomial and radial basis function (RBF) are used. Four basic kernels give an equation (6-11).

$$\begin{aligned}
 \text{linear: } K(x_i, x_j) &= x_i^T x_j \\
 \text{polynomial: } K(x_i, x_j) &= (\gamma x_i^T x_j + r)^d, \quad \gamma > 0 \\
 \text{RBF: } K(x_i, x_j) &= \exp(-\gamma \|x_i - x_j\|^2), \quad \gamma > 0 \\
 \text{sigmoid: } K(x_i, x_j) &= \tanh(\gamma x_i^T x_j + r)
 \end{aligned} \tag{6-11}$$

γ , r , and d are kernel parameters.

Initially SVM is designed to classify two classes; it is extended to multi class problems. SVM is used to classify ECG arrhythmia by (Homaeinezhad et al., 2012; Shen et al., 2010; Y. Wu & Zhang, 2011). In this study, the Gaussian radial basis function (Keerthi & Lin, 2003) defined in equation (6-11) is used, for classification of eight ECG beats with feature set developed from independent components and dynamic features. The performance of both ANN and SVM are nearly same. Due to form of function learned by ANN and SVM is typically the same. A single hidden layer neural network uses exactly the same form of the model as an SVM.

Chapter 7 Proposed Method- Hypothesis

The purpose of this chapter is to present the proposed methods, theory and related information about the proposed algorithms for noise removal and classification. The subsequent chapters are for better understanding of the work.

7.1 Motion Artefacts Removal: Proposed Method

This study is to exploit the independent component analysis for separation of the motion related artefacts blended within the ECG signals utilizing the independent statistical features of the ECG signals and motion related noise. It is followed by artificial neural network (ANN) classifier to work on the “cleaned” ECGs to categorize them into normal and abnormal patterns.

Motion related artefacts and EMG interferences are generally considered as the most troublesome, since they can mimic the appearance of ectopic beats and therefore cannot be removed by straightforward filtering (Goldberger et al., 2000; Moody et al., 1984). Due to body movements, ECG signals are contaminated or even corrupted by motion artefacts also known as ‘em’ artefacts. To improve the robustness of pattern recognition and classification of ambulatory ECGs, elimination of these ‘em’ artefacts is the key.

7.2 Proposed Algorithm

In this study a novel method is proposed to place an extra electrode on the body where ECG is minimal, thus the signals from ECG lead II will have a part of the noise, N' (which is a linear attenuated version of the motion related noise source N) and the signals from the extra leads, where the ECG is minimal, will contain mostly the noise N'' . Since noise appears in both

leads are homogeneous, all resulted from motions, the signals from the two leads are representative of the ECG and the noise with different mixing ratios. The ICA can be used to separate out the pure ECG and the motion related noise.

ICA based Blind source separation techniques could be used for separating ECG and ambulatory noise, as these signals are uncorrelated (Castells, Rieta, Millet, & Zarzoso, 2005; Hyvärinen & Oja, 2000a). For ICA blind source separation, multi-lead ECG recording is required and the ECG and noise to be removed should be independent of each other. To investigate the application of ICA for separation of electrode motion artefact using just two signals, ECG from modified lead II are taken from the MIT-BIH dataset and electrode motion noise signal from MIT noise stress test database.

$$\begin{aligned} \textit{Modified Lead II} &= S1 + N' \\ \textit{Limb Electrode} &= S2 + N'' \end{aligned} \quad (7-1)$$

Where modified lead II records ECG signal and part of electrode motion noise, the limb electrode also measures ECG signal and electrode motion noise, but ECG is minimal in this part so the majority of limb electrode signal is composed of 'em' noise. There are other parts as well, where the ECG is minimal like lumbar curve, which can also be used for motion artefact recording.

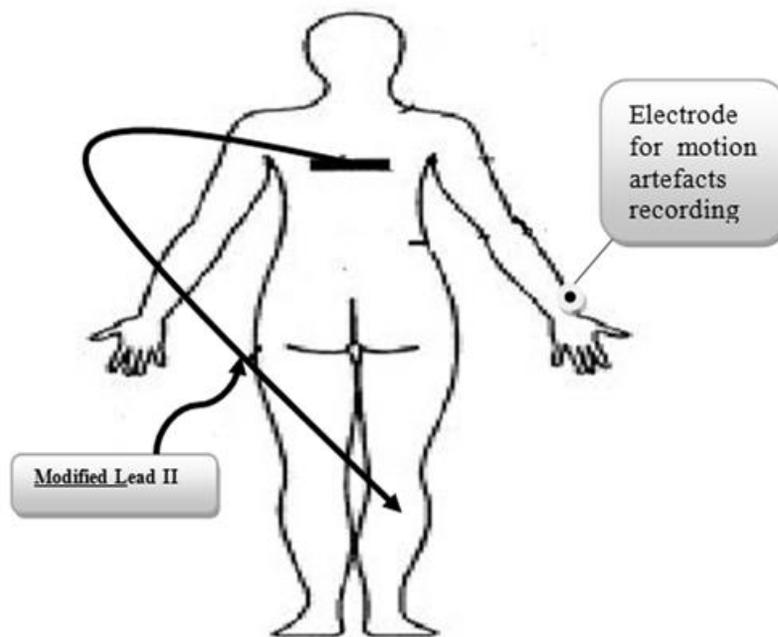


Figure 7-1 Proposed setup for removing motion artefacts

FastICA algorithm is used to separate the pure ECG and Motion artefacts. The selection of FastICA algorithm is based on comparative study of different ICA algorithms in the previous work of the same author (Sarfranz, Li, & Javed, 2011). The proposed setup adopted for this study is shown in figure 7-1.

With two independent components separated out from the ICA algorithm, one has to determine which one is the ECG. Visual inspection is certainly not desirable. In practice, the separated components tend to have more distinctive properties than the original signals both in time and frequency domains. Hence the statistical properties of these waveforms are employed and automatic recognition is done using kurtosis. The kurtosis is the fourth-order cumulant as explained in section 4.5 . The kurtosis is zero for Gaussian densities. For

continuous noise, the Kurtosis value is much smaller compared with that of normal ECG. In this work, a threshold of 30 is selected, a component whose modulus of kurtosis is below this threshold will be considered as continuous noise. When no component has a kurtosis over a specified threshold, then the component with the maximum kurtosis is selected. The main reason for choosing kurtosis is its simplicity. Computationally, kurtosis can be estimated by using the fourth moment of the sample data. The correlation coefficient is also used to differentiate noise and clean ECG. The value of 0.2 is set for correlation coefficient any value more than this is taken as noise and below that it is clean ECG.

The flowchart depicting procedure followed for removing motion artefact is shown in figure 7-2.

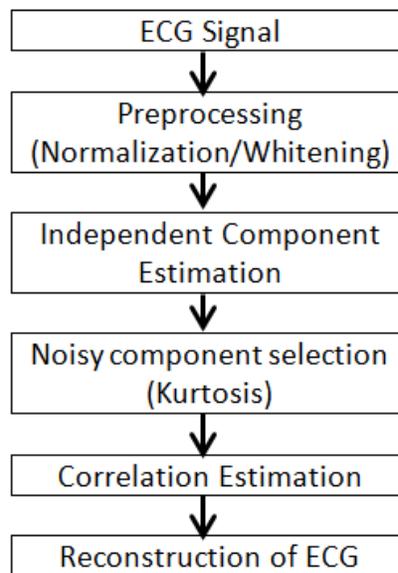


Figure 7-2 Motion artefact removal flowchart

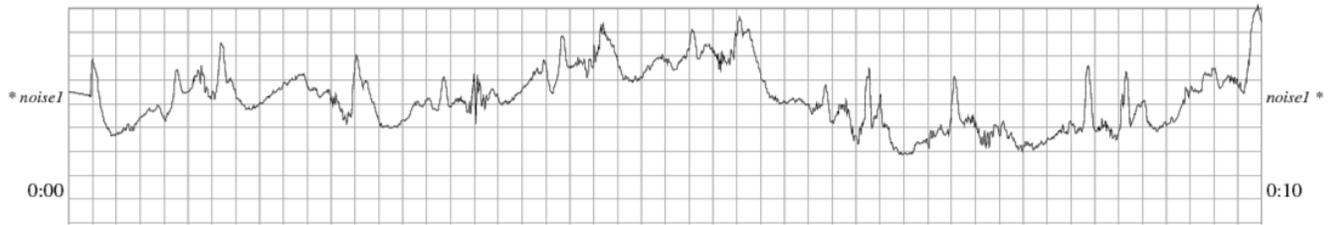


Figure 7-3 Sample of 'em' noise 10 Sec
 Picture adapted from (Goldberger et al., 2000)

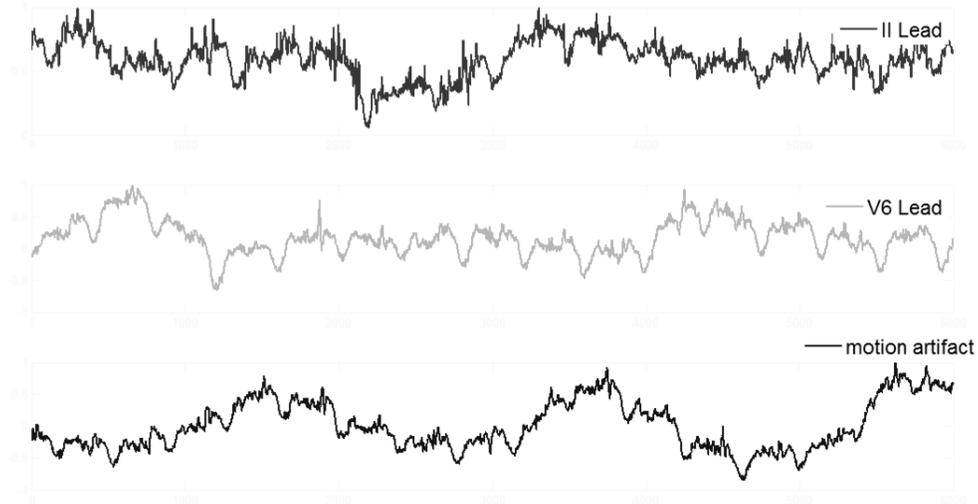


Figure 7-4 Ambulatory ECG with motion artefact

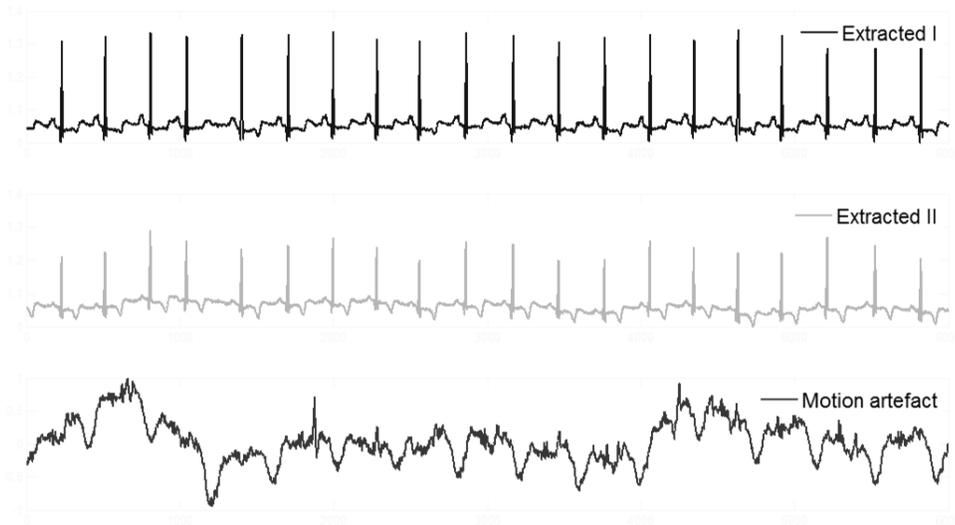


Figure 7-5 Extracted signal from two leads

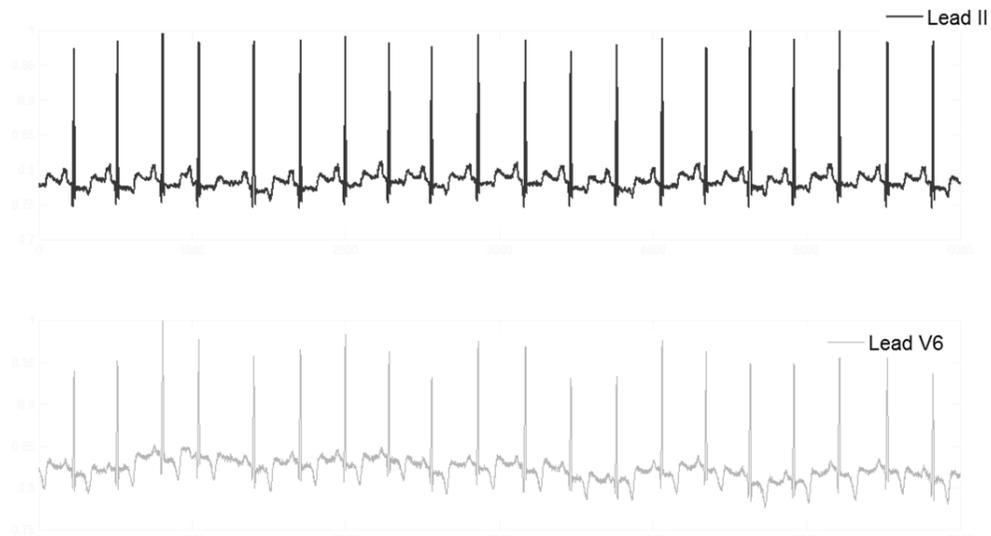


Figure 7-6 Original ECG data

To explain the noise overlapping and its effect on ECG signal, a motion artefact signal is mixed randomly on two lead ECG as shown in Figure 7-4 Ambulatory ECG with motion artefact. Both the ECG leads have lost maximum morphological information and manual inspection cannot reveal any meaningful information on current state. FastICA algorithm is

applied to two lead ECG with noise and it is able to separate the noise from ECG as shown in Figure 7-5 Extracted signal from two leads. When the extracted signal and original signal in Figure 7-6 Original ECG data are compared, it is clearly visible that ICA can successfully recover clean ECG from noisy mixture.

In order to evaluate a motion artefact reduction algorithm a database is generated by combining the clean ECG signal with different level of noise. For each type of arrhythmias 100 sets of 10 seconds are selected from the data set. 6 combination of different SNR value ranging from 24 dB to -12 dB are obtained for each data set. In total 4800 different combinations of ECGs with different SNR are obtained. The sampling frequency of ECGs is 360 Hz. Clean ECG and noise signal are then combined to get a signal with desired SNR value. This is achieved by multiplying 'em' signal by a gain factor and adding it to clean ECG. SNR values ranging from 24 to -12 dB instep of six are used. The output signals are then compared to reference 'em' signal. Signal quality improvement is measured by performance of beat detection algorithm and by the classification performance of noisy and clean signal.

7.3 Classification with Feature extraction from ECG using

Electrocardiogram (ECG) reflects the activities of the human heart and reveals hidden information on its structure and behaviour. The information is extracted to gain insight that assists to explain and differentiate diverse pathological conditions. This is traditionally done by an expert through visual inspection of ECGs. The complexity and tediousness of this onus hinder long-term monitoring and rapid diagnoses. Therefore computerised and automated ECG signal processing is sought after. In this study an algorithm that uses independent component analysis (ICA) to improve the performance of ECG pattern recognition is proposed. The algorithm deploys the basis functions obtained via the ICA of typical ECG to extract ICA features of ECG signals for further pattern recognition, with the hypothesis that components of an ECG signal generated by different parts of the heart during normal and arrhythmic cardiac cycles might be independent. The features obtained via the ICA together with the R-R interval and QRS segment power are jointly used as the input to a machine learning classifier, an artificial neural network in this case. Results from training and validation of the MIT-BIH Arrhythmia database show significantly improved performance in terms of recognition accuracy. This new method also allows for the reduction of the number of inputs to the classifier, simplifying the system and increasing the real-time performance. This study presents the algorithm, discusses the principle algorithm and presents the validation results.

7.4 Proposed Method

The basic ICA model is that a vector of unknown sources \mathbf{s} is not observed directly, but from a linear combination of them. $\mathbf{x} = \mathbf{A}\mathbf{s}$, where \mathbf{x} is a $N \times 1$ column vector, \mathbf{A} is a $N \times N$ mixing matrix, and \mathbf{s} is a $N \times 1$ column vector of the source signals. The columns of \mathbf{A} $\{\mathbf{a}_i\}$ are used as the basis functions which generate the observed signals. The inverse of \mathbf{A} , $\mathbf{W} = \mathbf{A}^{-1}$, transforms the original signals into the unknown source coefficients, $\mathbf{s} = \mathbf{W}\mathbf{x}$.

The objective of ICA algorithm is to find out the basis function by adapting and learning from the ECG data \mathbf{x} in this study. The cost function in ICA can be minimization of mutual information of demixing model output as used in Infomax algorithm (Bell & Sejnowski, 1995) or the maximization of nongaussianity used in FastICA algorithms (Hyvarinen, 1999).

Using ICA as a feature extraction method, the basis functions \mathbf{a}_i is considered as the basis features and the component of \mathbf{s} , \mathbf{s}_i is considered to be the coefficient for each basis feature in the feature space. So the model can be written as $\mathbf{x} = \sum_{i=1}^N \mathbf{s}_i \mathbf{a}_i$. As compared to correlation-based transformations such as principal component analysis (PCA), ICA not only uses second-order statistics to decorrelate the signals, but also uses high-order statistics to reduce high-order dependencies. Thus the output coefficients corresponding to different basis functions are as statistically independent as possible.

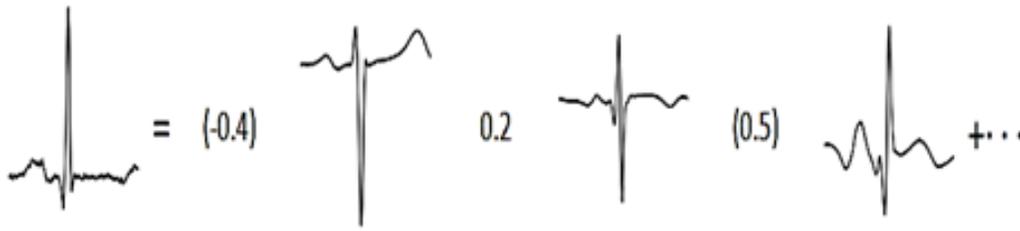


Figure 7-7 A normal heartbeat is a linear combination of ICA bases with Coefficients,
 e.g. $(-0.4) \cdot a_i + (0.2) \cdot a_j + 0.5 \cdot a_k + \dots$

ICA basis functions $\{a_i, a_j, a_k, \dots\}$ is scaled by the corresponding coefficients $\{s_i, s_j, s_k, \dots\}$. The heartbeat segment is represented mostly by $\{a_i, a_j, a_k\}$, as most of the other coefficients are almost zero. The ICA basis functions reveal the statistical structures of the single heartbeat segments of ECG as shown in figure 7-7.

The ECG signals are obtained from the MIT-BIH arrhythmia database for study. Since most of the diagnostic information of the ECG signal lies around R peak i.e. QRS segment, which is about 0.06 – 0.10 s, so designated portion on both sides of R peak is chosen. In this study 200 data points in each sample which is nearly 0.0556 s of ECG signals are selected. The sampling frequency of this signal is 360 Hz. The extracted sample of 200 data point has all the required information of ECG pulse including P and T wave as well, which gives us complete information contained in single pulse including the cases of noise presence along with the original signal. Figure 7-8 gives the general idea of how different arrhythmias have different beat characteristics. Same signal size and specification is used in our earlier work (Sarfranz et al., 2011; Sarfranz, Li, & Javed, 2013; Sarfranz & Li, 2013).

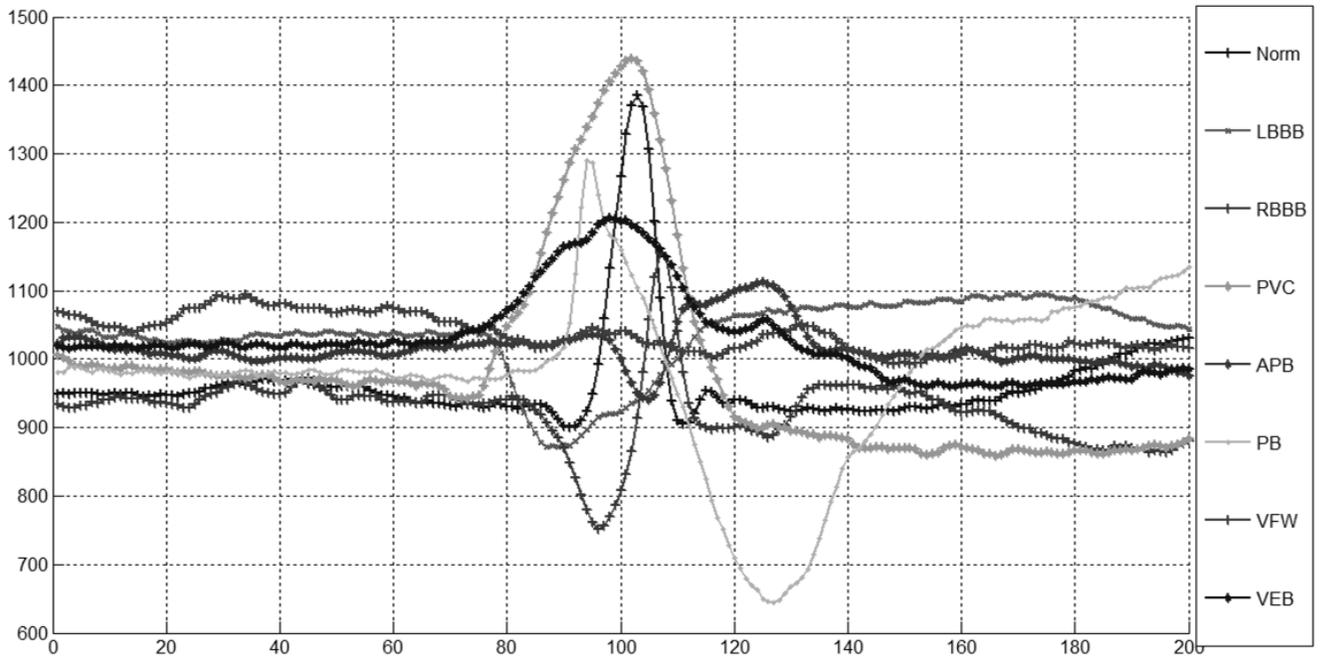


Figure 7-8 ECG heartbeats of eight types

It is better to define the statistics of the classification results used in this study as follows

TP: True positive, correctly classified abnormal beats.

TN: True negative, correctly classified normal beat.

FP: False positive, incorrectly classified normal beats.

FN: False negative, incorrectly classified abnormal beats.

Sensitivity is defined as a measure to test positive samples.

$$SE = \frac{TP}{TP + FN} \quad (7-2)$$

Specificity is defined as a measure to test negative samples.

$$SP = \frac{TN}{TN + FP} \quad (7-3)$$

Accuracy is defined as a measure to test samples correctly.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7-4)$$

Positive Predictivity or Positive Predictive Values (PPV) is defined as ratio of true positive with total number of positive cases.

$$PPV = \frac{TP}{TP + FP} \quad (7-5)$$

Type	MIT-BIH data, file reference
NORM	100,101,103,105,108,112,113,114,115,117,121,122,123,202,205 219,230,234
LBBB	109,111,207,214
RBBB	118,124,212,231
PVC	106,119,200,203,208, 213,221,228,233,116,201,210,215
APB	209,222,232,220,223
PB	102,104,107,217
VFW	207
VEB	207

Table 7-1 Record and number of ECG sample used

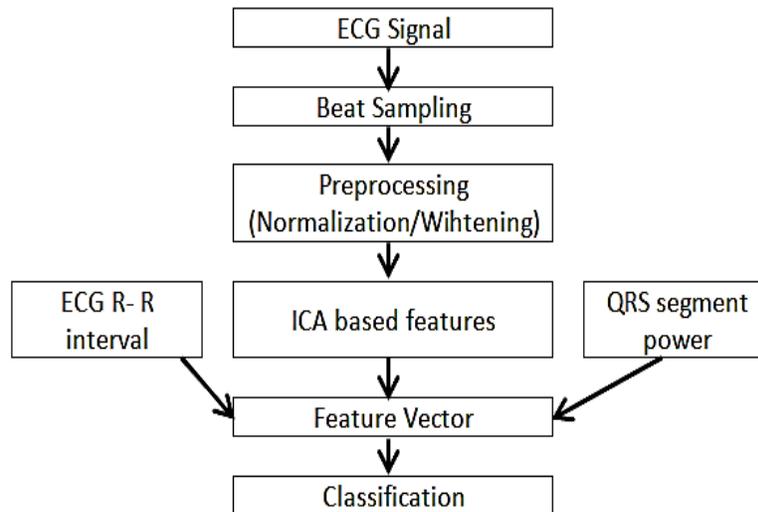


Figure 7-9 Block diagram of the proposed feature extraction and classification system

Flowchart to depict the procedure for motion artefact removal is shown in figure 7.2 after finding the unmixing matrix W using FastICA, and then reordering the row of the W matrix by the norm of the row from the smallest to the largest. The independent components (ICs) are calculated. They are employed as bases for calculation of feature vector.

It is interesting to note fact that arrhythmias are different from the normal heart in terms of both morphology and dynamics. Two dynamic features are introduced to describe the rhythm of a heartbeat, namely pre-RR interval and post- RR interval. The pre-RR-interval is the RR interval between a given heartbeat and the previous heartbeat. The post-RR-interval is the RR-interval between a given heartbeat and the following heartbeat. Both of them are calculated using pan Tompkins algorithm for QRS detection (Pan & Tompkins, 1985). In both training and testing stage, each sample of the training set is projected onto the ICs. The

projections constitute the ICA-feature vector. Along with the RR interval, and QRS signal power the feature vector is obtained. Flowchart to depict the procedure for feature extraction and classification is shown in figure 7.2 .

In this study BPNN is used as classifier in the training stage. The input is feature vectors and the output value is set at 1, 2, 3, to 8 for NORM, LBBB, RBBB, or PVC (8 types of ECG). The aim of this stage is for adjusting the best parameters in neural networks for the best classification. To study the effect of the numbers of ICs in ECG beat classification, the numbers of ICs vary from 5 to 40 and their effects is investigated. The results presented in this study achieved the best performance of classifier with 15 ICs. 10 fold approach is used to select the number of ICs. Further increasing the number of ICs from 15 do not improve classification accuracy regularly. Even though some improvement is recorded, but it is neglected, due to non-uniformity. After the neural network is trained, it is applied to the other set of ECG samples for testing the performance of the classifier and calculating the specificity and sensitivity.

Classification is done using Back Propagation neural network (BPNN) implemented in MATLAB software. ANN are widely used classifier for ECG (Al-Fahoum & Howitt, 1999; Belgacem, Chikh, & Reguig, 2003a; Jiang, Zhang, Zhao, & Albayrak, 2006a; Wang, Chiang, Yang, & Hsu, 2012; S.-N. Yu & Chou, 2007). Back-propagation neural network (BPNN) used in this study is a three-layer feed-forward structure (Jang, Sun, & Mizutani, 1997a). The first layer is the input layer that has the ICA features , pre-RR-interval, post-RR-interval and QRS segment power as inputs.

In this study 15 ICs are selected to be part of feature vectors. The average accuracy, increases rapidly at small numbers of ICs and then reaches a plateau at around 15 ICs, at even higher IC numbers, the average accuracies stay at around 99.4 %. Further increase in IC number does not significantly increase the accuracy of the classifier. 18 neurons are used in the input layer, related to morphological and dynamic features. The second layer, also called the hidden layer, has 20 neurons and the output layer has 8 neurons for classification, which is equal to the number of ECG beat types to be classified. In this study, the hyperbolic tangent functions are used in the first and second layer, and the identity function is used in the output layer. The weight and bias values in the BPNN are updated by Levenberg-Marquardt optimization method (Jang, Sun, & Mizutani, 1997b) with a learning rate of 0.1. A criterion of 0.01 in mean-square-error is empirically determined to terminate the iterations in the training phase of the classifier. Time taken for the training of classification is about 1.6 seconds in the Matlab computing environment based on the average over 10 times.

Chapter 8 Results and Discussion

The entire experiments was simulated using MATLAB. The MIT-BIH arrhythmia and noise-stress database were used. An annotated and validated database is very important for the study of ECG signal processing and pattern recognition in general, and such a “standard” database is particularly useful in this study. This allows for the validation of the newly developed algorithms and compared with the results from others work. The selection of MIT-BIH database is natural choice because it is completely annotated by medical specialists and arguably the most popular one used by many other authors and quoted in numerous important publications in this field e.g. (Belgacem, Chikh, & Reguig, 2003b; Chou & Yu, 2008b; Jiang et al., 2006 a; Moody, Mark, & Goldberger, 2001; Wang et al., 2012). The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings. A total of 9800 sample segments attributing to eight ECG beat types are selected from the MIT-BIH arrhythmia database for experiments. The eight beat types used in the study are normal beat (NORM), left bundle branch block beat (LBBB), right bundle branch block beat (RBBB), atrial premature beat (APB), premature ventricular contraction (PVC), paced beat (PB), ventricular flutter wave (VFW), and ventricular escape beat (VEB). The types and numbers of the ECG beat exploited in the study are summarized in Table 7-1 Record and number of ECG sample used, out of 9800 beats half of the ECG beats are used for training and the other half for testing the classifiers performance. Though the dataset is large enough but some arrhythmias are present in few patients which results in small dataset.

The MIT-BIH Noise Stress Test Database includes typical ambulatory noise recordings made using physically active volunteers. Standard ECG recorders, leads, and electrodes are used; the electrodes are placed on the limbs in positions where the subject’s ECGs are virtually

invisible. Electrode motion artefact is generally considered the most troublesome, since it can mimic the appearance of ectopic beats and cannot be removed easily by simple filters, as can noise of other types by (Moody, Muldrow, & Mark, 1984). The ECG recordings are created from the MIT-BIH Arrhythmia Database, to which calibrated amounts of noise from record 'em' are added. Six set of noisy ECG signals are developed with SNR ranging from 24 dB to -6 dB, with interval of six as used in MIT-NSTDB. Matlab software is used for the signal processing, pattern recognition, visualization and user interface in this study. The ICA BSS algorithm used is the FastICA which is experimentally verified to have the best result of noise removal for similar application (Sarfraz et al., 2011).

In order to evaluate a motion artefact reduction algorithm a database is generated by combining the clean ECG signal with different level of noise. For each type of arrhythmias 100 sets of 10 seconds are selected from the data set. 6 combination of different SNR value ranging from 24 dB to -12 dB are obtained for each data set. In total 4800 different combinations of ECGs with different SNR are obtained. The sampling frequency of ECGs is 360 Hz. Clean ECG and noise signal are then combined to get a signal with desired SNR value. This is achieved by multiplying 'em' signal by a gain factor and adding it to clean ECG. SNR values ranging from 24 to -12 dB instep of six are used. Figure 8-1 shows one example of a clean ECG; a pure noise and a combination of both signals. For evaluation the simulated signals with added 'em' noise are used as input to denoising algorithm. The output signals are then compared to reference 'em' signal. Signal quality improvement is measured by performance of beat detection algorithm and by the classification performance of noisy and clean signal.

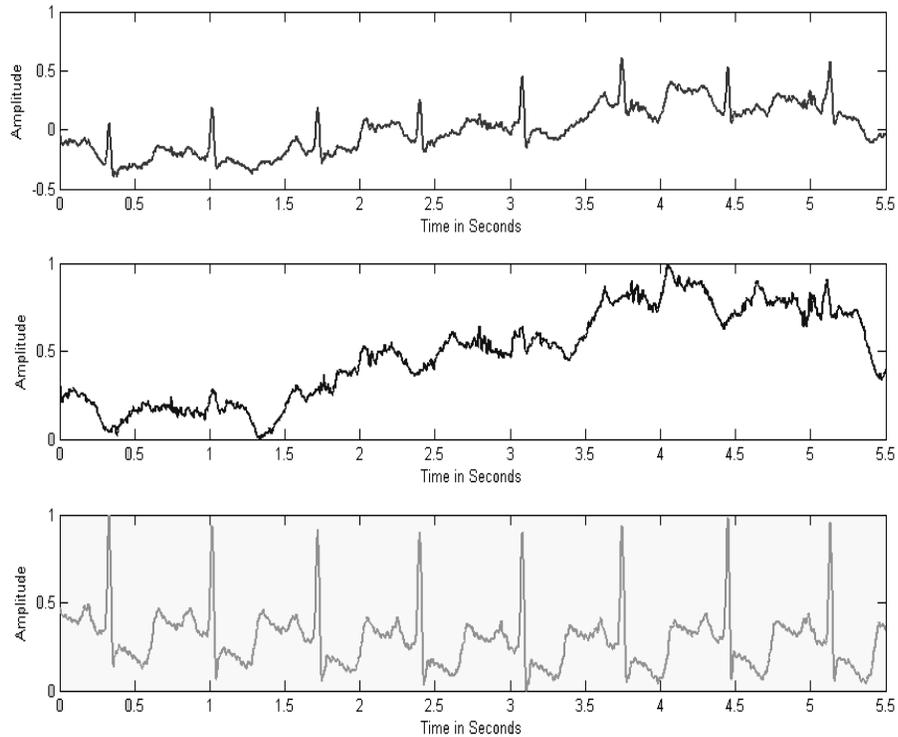


Figure 8-1 ECG Combine with noise (Top Panel), 'em' Noise middle panel, Extracted ECG with ICA bottom panel. Y Axis is normalized after ICA

Figure 8-2 shows an example of a clean ECG (before adding noise), a noisy ECG when noise is added to a SNR = -6 dB) and filtered ECG using ICA. The beat detection had a significant improvement after filtering as compared to the noise signal.

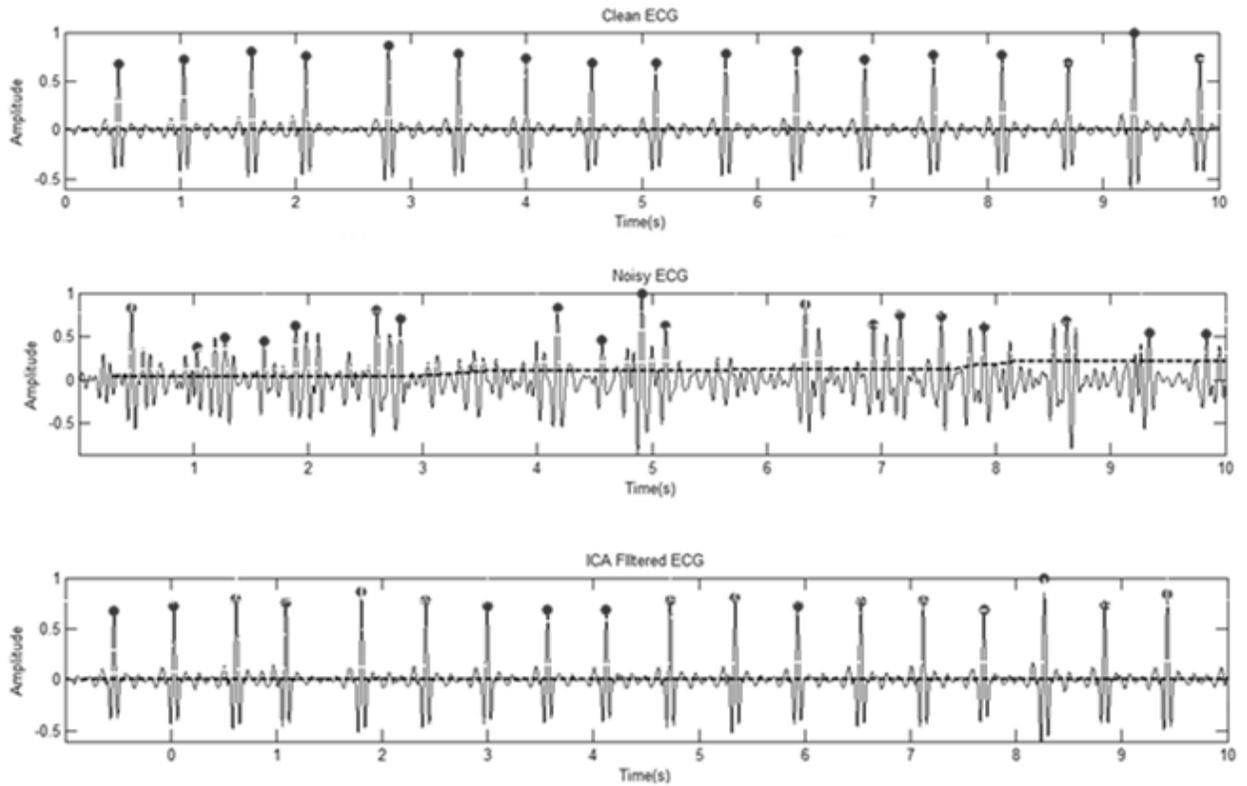


Figure 8-2 Pan Tomkins algorithm for beat detection on a sample of a clean ECG, a noisy ECG (SNR= -6 dB) & filtered ECG using ICA. The beat detection has significant improvement after ICA filtering.

To test the effectiveness of the ICA in the removal of motion related artefacts, ANN classifiers for normal and abnormal ECGs with ICA filtering and with basic filtering are tested with ECGs mixed with motion related noise in a multiple signal to noise ratios ranging from 24 dB to -12 dB are used to validate the proposed method. The sensitivity of a test is defined as the proportion of people with arrhythmia who will have a positive result. The PPV of a test is the proportion of people with normal ECG who will have a negative result. The higher value of positive predictive value and sensitivity indicate better classification and small error. Comparison results are shown in. It has clearly shown that when the signal to noise ratio increases, the performance of ANN classification without removing noise with ICA

algorithm decreases significantly with low recognition accuracy. With the ICA de-noising recognition accuracy improves significantly in poor signal to noise ration conditions. It is observed from the results that ICA noise removal improves the accuracy of the classification, for all levels of noise, but more significantly in the noisy ECG there is an improvement up to 40% in sensitivity and positive predictivity as shown in Figure 8-3 and Figure 8-4. The classification accuracy also records major improvement in noisy signals as shown in Figure 8-5.

Noise (dB) \ Sensitivity (%)	-12 dB	-6 dB	0 dB	6 dB	12 dB	18 dB	24 dB
ICA Filtering	93.3	94	95.1	97.1	97.6	98	97.9
Basic Filtering	55.6	61.1	80.6	83.2	92.9	96.5	96.5

Table 8-1 Comparison of classification Sensitivity before and after source separation.

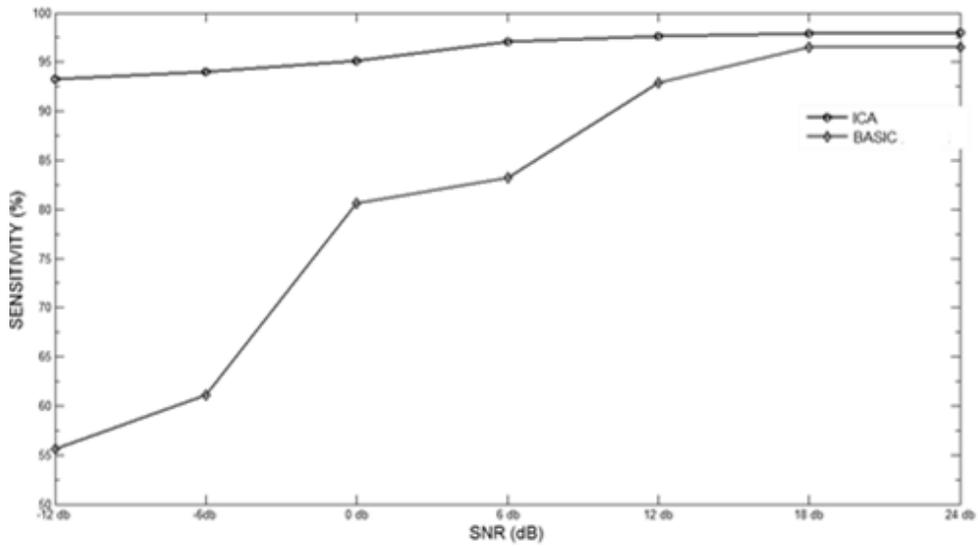


Figure 8-3 Sensitivity with ICA filtering and without filtering

Noise (dB) \ Predictivity (%)	-12 dB	-6 dB	0 dB	6 dB	12 dB	18 dB	24 dB
ICA Filtering	94.1	93.8	94.7	97.4	97.8	98.1	98.1
No Filtering	55.9	59.9	81.6	82.5	93.6	96.4	96.1

Table 8-2 Comparison of classification positive predictive value before and after source separation.

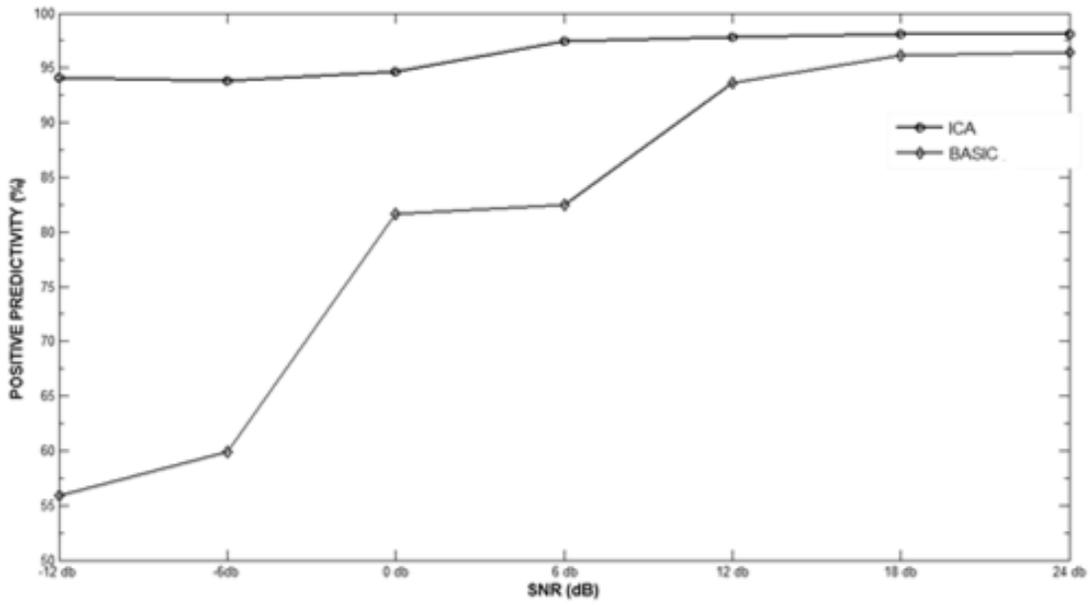


Figure 8-4 Positive Predictivity with ICA filtering and without filtering

Noise (dB) \ Accuracy (%)	-12 dB	-6 dB	0 dB	6 dB	12 dB	18 dB	24 dB
ICA Filtering	93.8	93.9	94.9	97.2	97.7	98.1	98
No Filtering	55.9	60.1	81.2	82.8	93.3	96.4	96.3

Table 8-3 Comparison of classification accuracy before and after source separation.

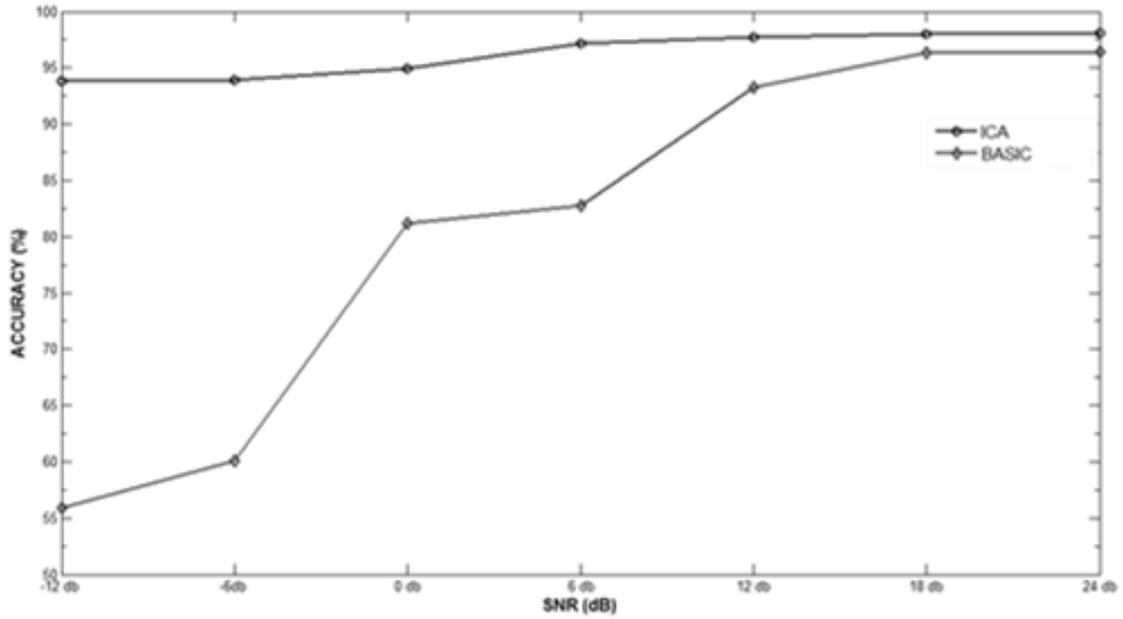


Figure 8-5 Accuracy with ICA filtering and without filtering

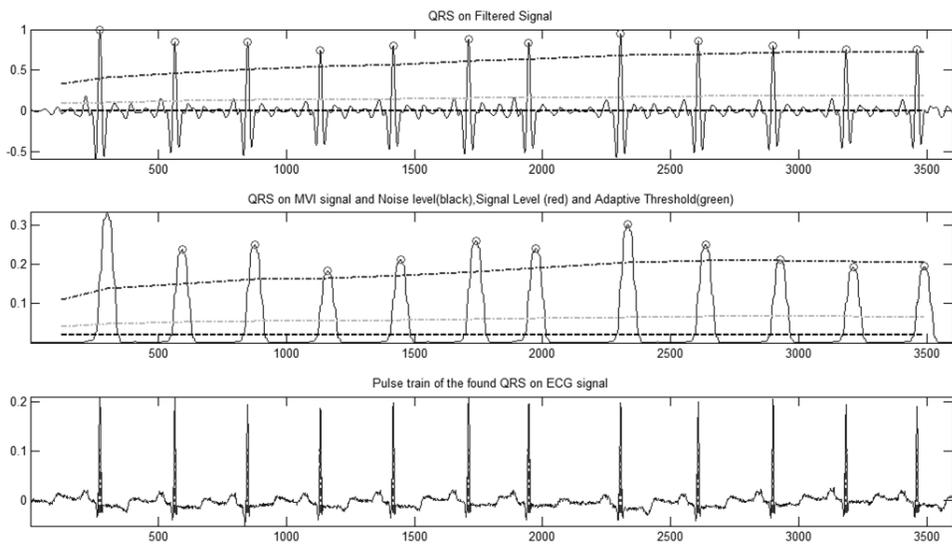


Figure 8-6 ECG beat detection using a Pan Tomkins algorithm

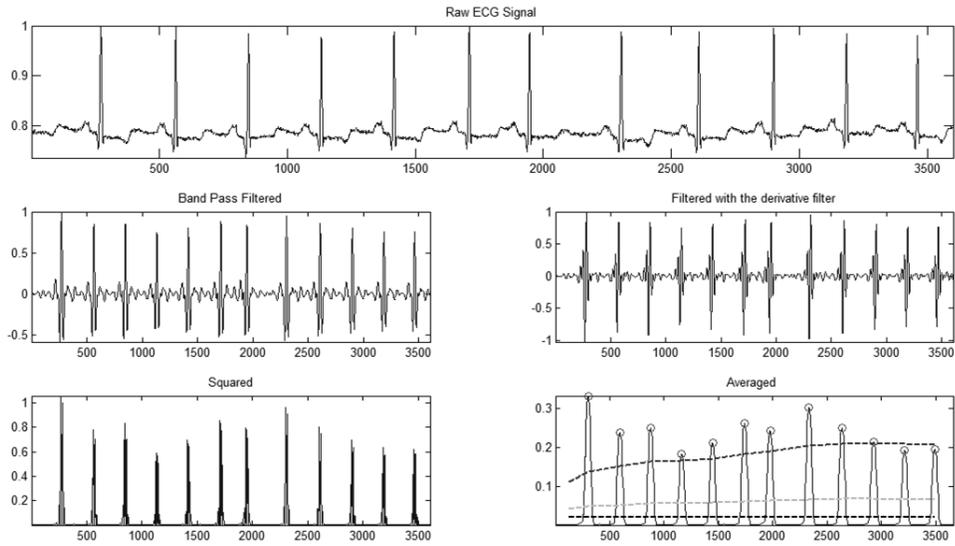


Figure 8-7 ECG QRS detection using a Pan Tomkins algorithm

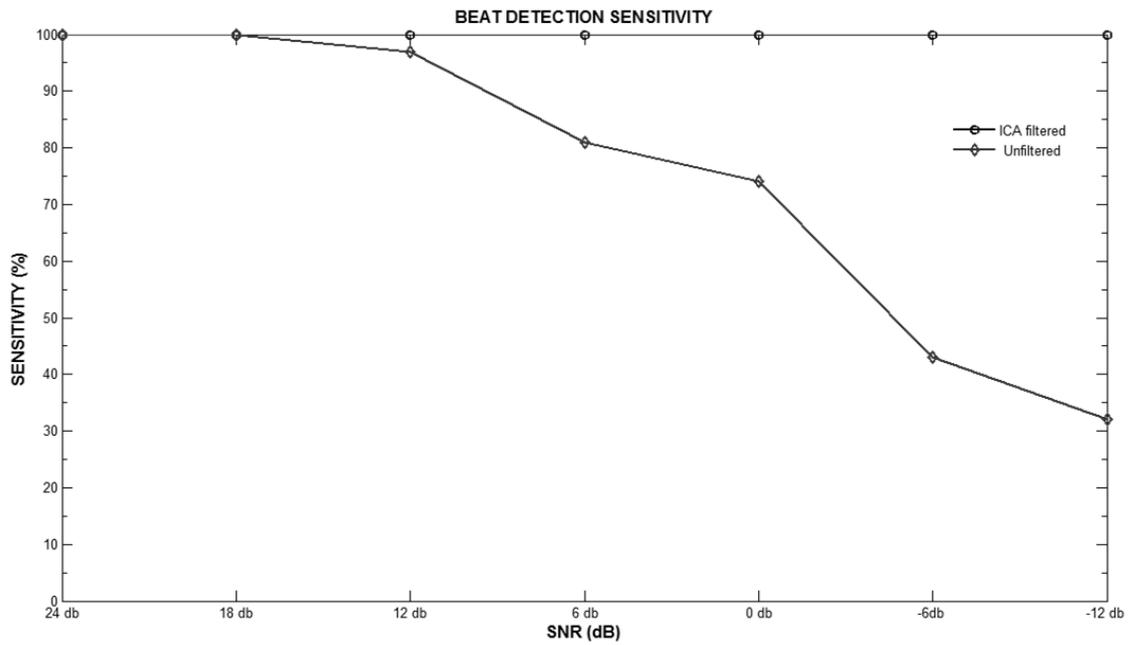


Figure 8-8 Beat detection comparison for LBBB arrhythmia with proposed ICA filtering and without filtering

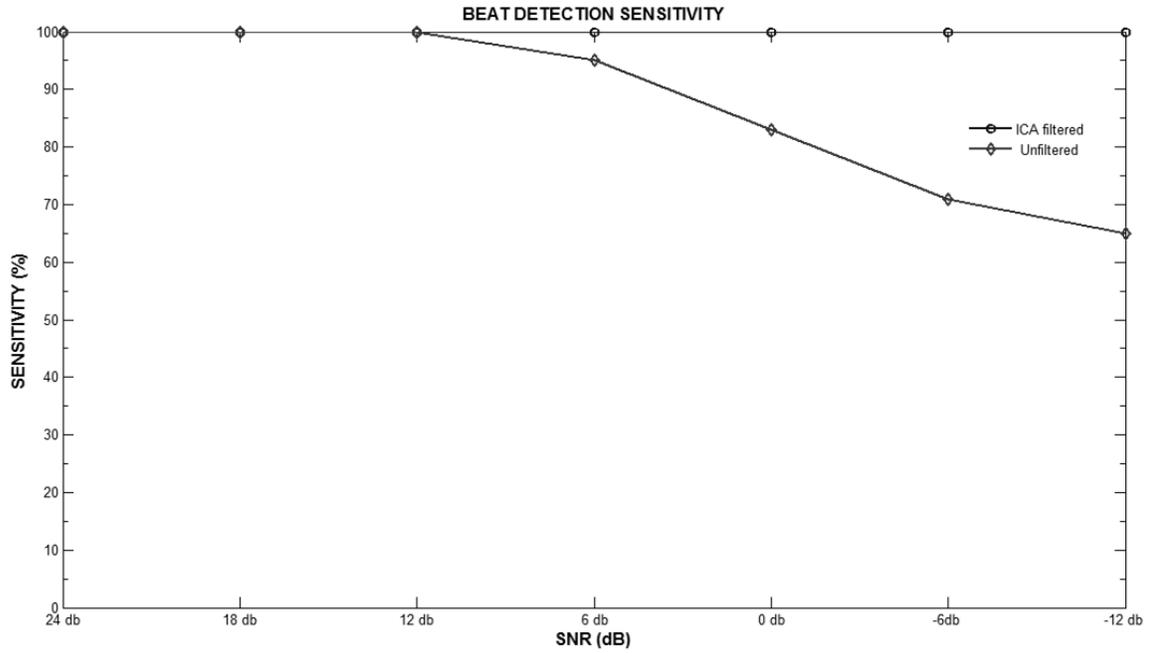


Figure 8-9 Beat detection comparison for Normal ECG beat with proposed ICA filtering and without filtering

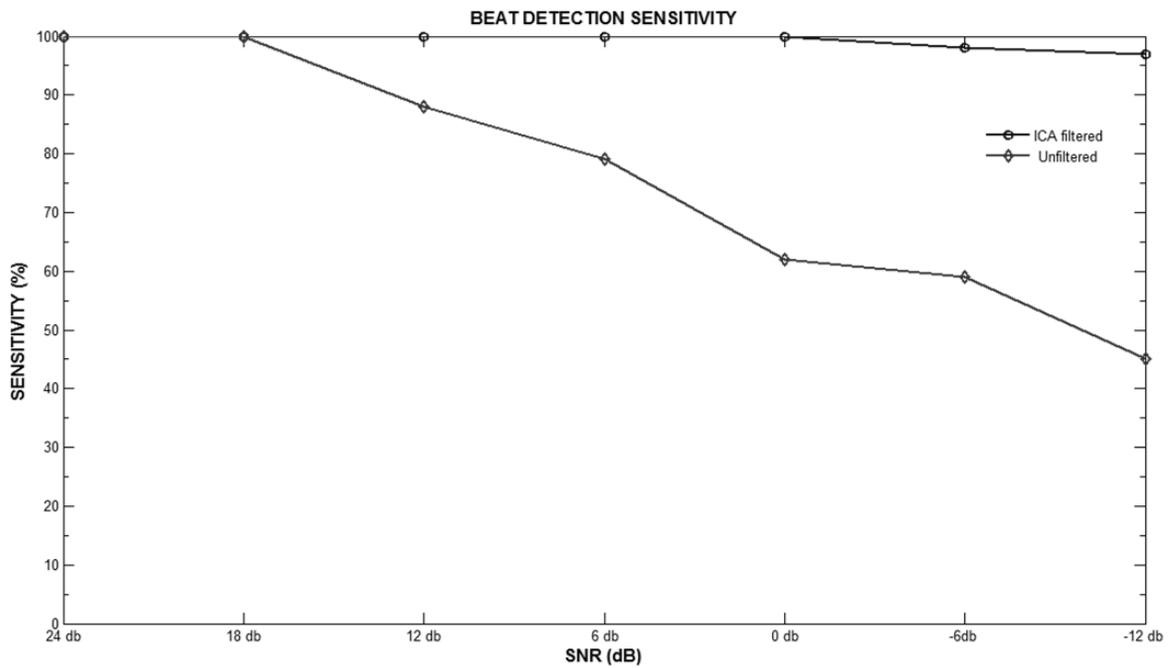


Figure 8-10 Beat detection comparison for RBBB ECG beat with proposed ICA filtering and without filtering

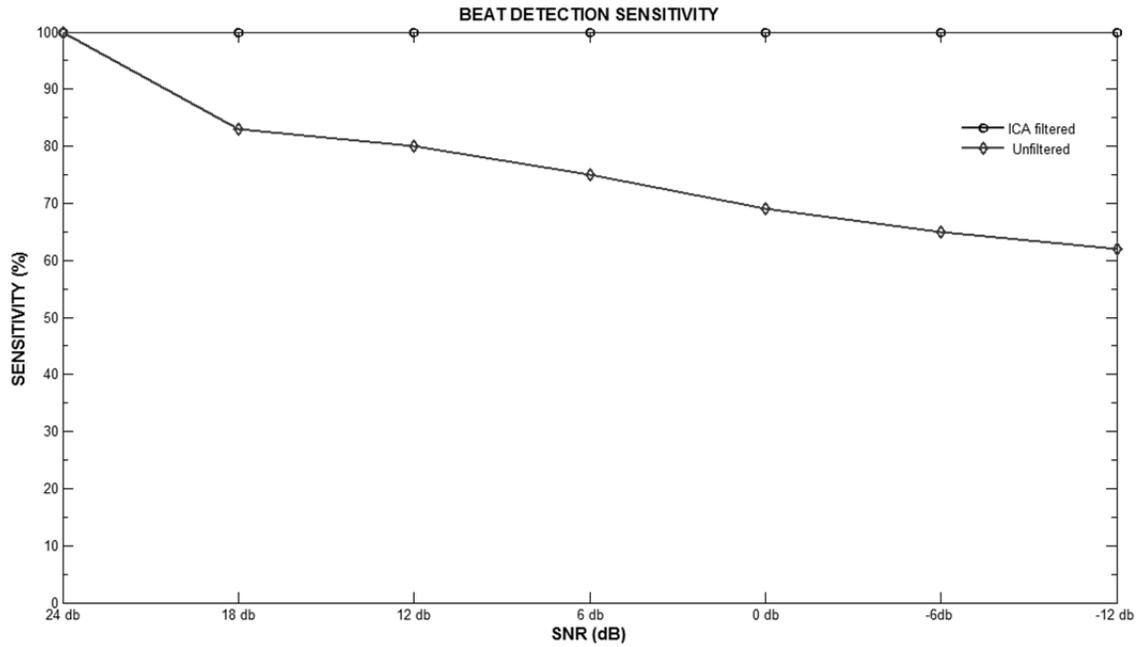


Figure 8-11 Beat detection comparison for VEB arrhythmia with proposed ICA filtering and without filtering

Significant performance enhancement is observed in beat detection in different arrhythmias with the proposed ICA filtering algorithm. It is also observed that different arrhythmias type have varied performance with beat detection due to different morphological features.

This section presents the results of classification of ECG arrhythmias into eight classes, using ICA for feature extraction.

All the experiments in this study are simulated using MATLAB. The ECG signal from MIT BIH arrhythmia database is subjected to QRS complex detection using Pan-Tompkins method. After QRS complex detection, 200 data point sample with 100 data points from the right of QRS peak, 99 data points to the left of QRS complex and the QRS peak itself is chosen as a segment of ECG beat. The choice of 200 samples around the R peak as a signal window length is such that it consists almost one cycle of cardiac activity. This duration is used in author's previous studies as well (Sarfranz et al., 2014; Sarfranz & Li, 2013).

Fifteen features of ICA and PCA along with three other features, pre RR interval, post RR interval and QRS segment power are fed to different classifiers, ANN and SVM for automated classification. The input layer consisted of 18 nodes corresponding to the 18 features used; a hidden layer of 20 neurons and an output layer of 8 neurons corresponding to the eight classes is used. The choice of 20 neurons in the hidden layer is by trial and error method. The three layer neural network with different number of hidden neurons is also tested. The highest accuracy is recorded with 20 neurons in the hidden layer. The ANN weights are updated using the error back propagation method of learning. In this study 10-fold cross validation technique is used for training and testing of the classifiers. The overall performance of the classifier is evaluated by taking the average of ten folds. The correct classification or misclassification is assessed as True Positive (TP), True Negative (TN), False Positive (FP) and False

Negative (FN). Based on these measures the sensitivity, specificity, positive predictive value and accuracy are determined.

Figure 8-13 shows accuracy in each of the ten folds of ANN classifier with the proposed features set. It can be observed from this figure, that very high overall accuracy is achieved with the proposed method. Figure 8-14 shows specificity in each of the ten folds for different arrhythmias. Normal beats having 100 % average specificity of 100 %, while VEB have 99.1 %. It can be observed from Figure 8-15 that ANN classifier with ICA extracted features set along with pre & post R- intervals and QRS segment power provides highest sensitivity for arrhythmias. 100 % sensitivity is recorded for nearly all arrhythmias types. The performance of the proposed classification method is tabulated in Table 8-4, Table 8-5 and Table 8-6.

1	200 13.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
2	0 0.0%	199 13.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
3	0 0.0%	0 0.0%	200 13.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
4	0 0.0%	1 0.1%	0 0.0%	198 13.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.5% 0.5%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	200 13.8%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	200 13.8%	0 0.0%	0 0.0%	100% 0.0%
7	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	200 13.8%	0 0.0%	99.0% 1.0%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	50 3.4%	100% 0.0%
	100% 0.0%	99.5% 0.5%	100% 0.0%	99.0% 1.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	99.8% 0.2%
	1	2	3	4	5	6	7	8	

Figure 8-12 Classification into eight classes with proposed method

Figure 8-12 shows the confusion matrix with sensitivity, specificity and overall accuracy of ECG beats classification to eight classes. Each row depicts one type of arrhythmia. The last column of matrix shows the individual positive predictive value of each arrhythmia type. While the last row of matrix shows the individual sensitivity of each arrhythmia type. Highest accuracy of 99.8% was reached as shown in the bottom right cell of the confusion matrix.

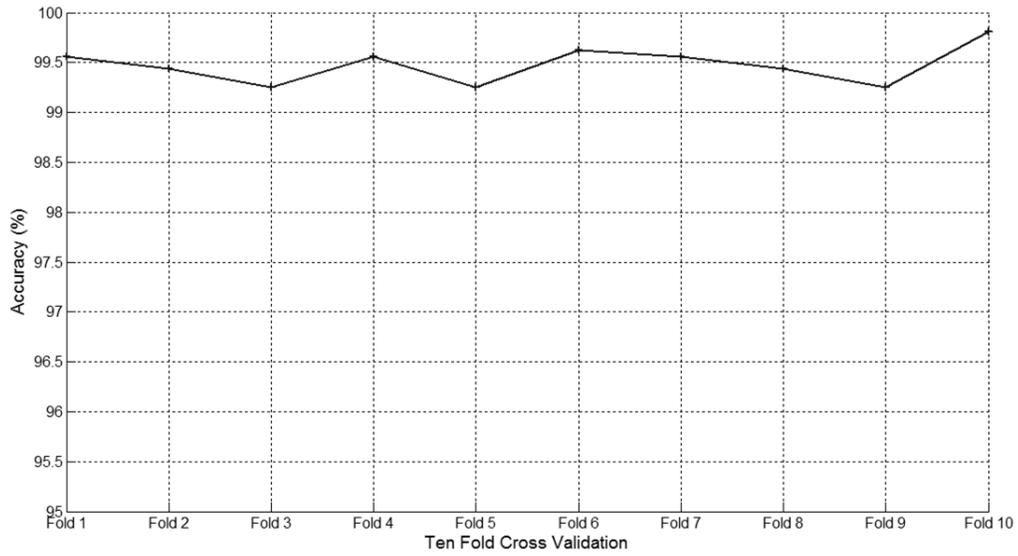


Figure 8-13 Results of Arrhythmia Classification Accuracy with proposed method

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
Accuracy	99.6	99.4	99.3	99.6	99.3	99.6	99.6	99.4	99.3	99.8

Table 8-4 Detailed accuracy for different arrhythmias with proposed method

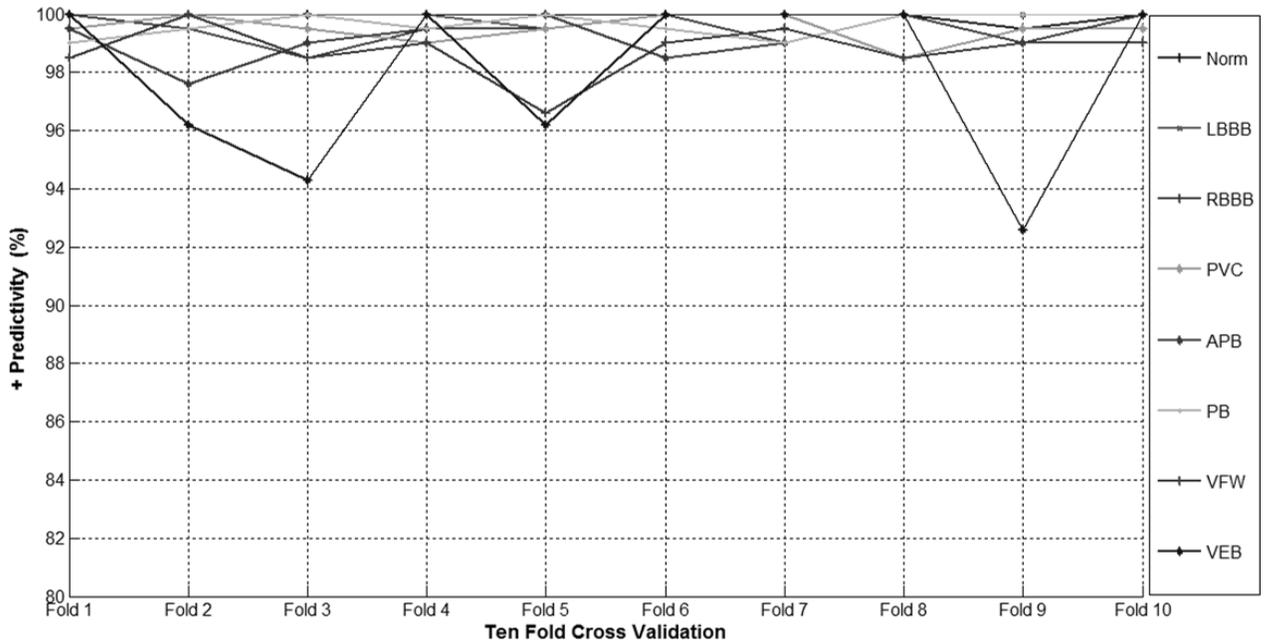


Figure 8-14 Results of positive predictive value with the proposed method

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold
NORM	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.5	100.0
LBBB	100.0	100.0	100.0	100.0	99.5	100.0	100.0	100.0	100.0	100.0
RBBB	100.0	99.5	98.5	99.5	99.5	100.0	99.0	100.0	99.0	100.0
PVC	99.5	100.0	99.5	99.0	99.5	100.0	100.0	98.5	99.5	99.5
APB	99.5	97.6	99.0	99.5	100.0	98.5	99.0	100.0	99.0	100.0
PB	99.0	99.5	100.0	99.5	100.0	99.5	99.0	100.0	100.0	100.0
VEB	98.5	100.0	98.5	99.0	96.6	99.0	99.5	98.5	99.0	99.0
VFW	100.0	96.2	94.3	100.0	96.2	100.0	100.0	100.0	92.6	100.0

Table 8-5 Detailed positive predictive value for different arrhythmias with proposed method

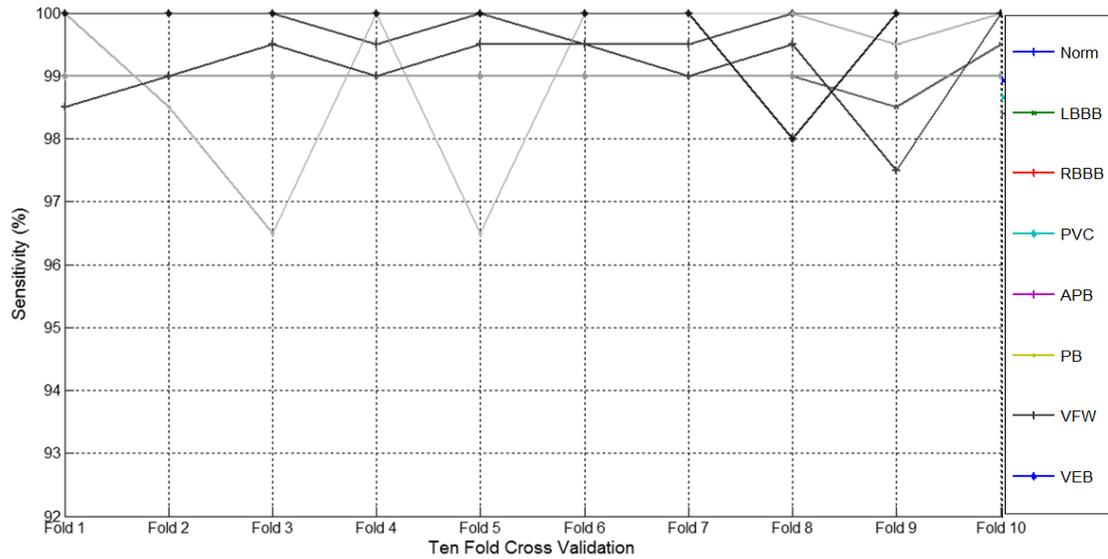


Figure 8-15 Results of sensitivity with the proposed method

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold
NORM	100	100	100	100	100	100	100	100	100	100
LBBB	99	99	99	99	99	99	99	99	98.5	99.5
RBBB	100	100	100	99.5	100	99.5	99.5	100	99.5	100
PVC	99	99	99	99	99	99	99	99	99	99
APB	100	100	100	100	100	100	100	100	100	100
PB	100	98.5	96.5	100	96.5	100	100	100	99.5	100
VEB	98.5	99	99.5	99	99.5	99.5	99	99.5	97.5	100
VFW	100	100	100	100	100	100	100	98	100	100

Table 8-6 Detailed sensitivity for different arrhythmias with the proposed method

Figure 8-16 shows the accuracy for different folds of SVM classifiers using proposed features set. Average accuracy of 99 % is recorded. Highest being 99.3 %. It can be noted from this figure, that SVM provided less accuracy, whereas ANN provided highest accuracy. Although this performance can be tested by changing different kernel function, but that is beyond the scope of this work and can be looked in future work for performance analysis of different classifiers with the proposed features set. The performance of specificity and sensitivity due to ICA features extracted different folds of SVM classifier is shown in Figure 8-17 and Figure 8-18. NORM beats showed highest specificity of nearly 100 %, while VEB and VFW beat recorded 98.6 % and 96.4 % specificity mainly due to a small data set for these two type beats.

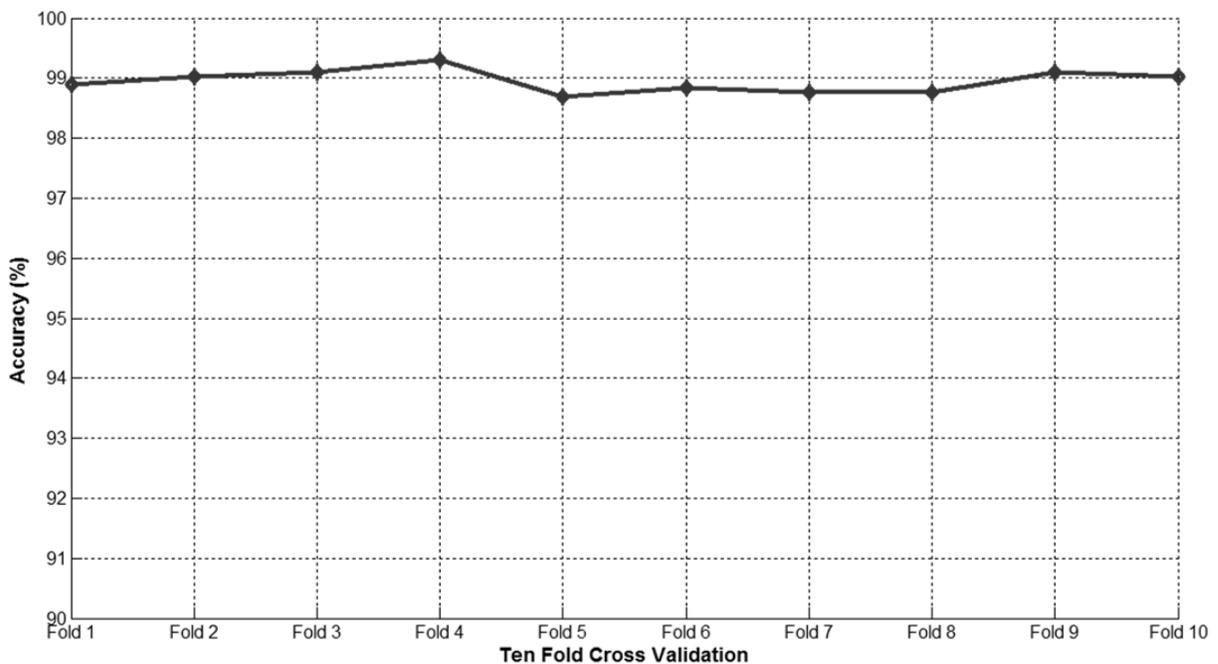


Figure 8-16 Results of Arrhythmia Classification Accuracy with SVM classifier

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
Accuracy	98.90	99.03	99.10	99.31	98.69	98.83	98.76	98.76	99.10	99.03

Table 8-7 Detailed accuracy for different arrhythmias with SVM classifier

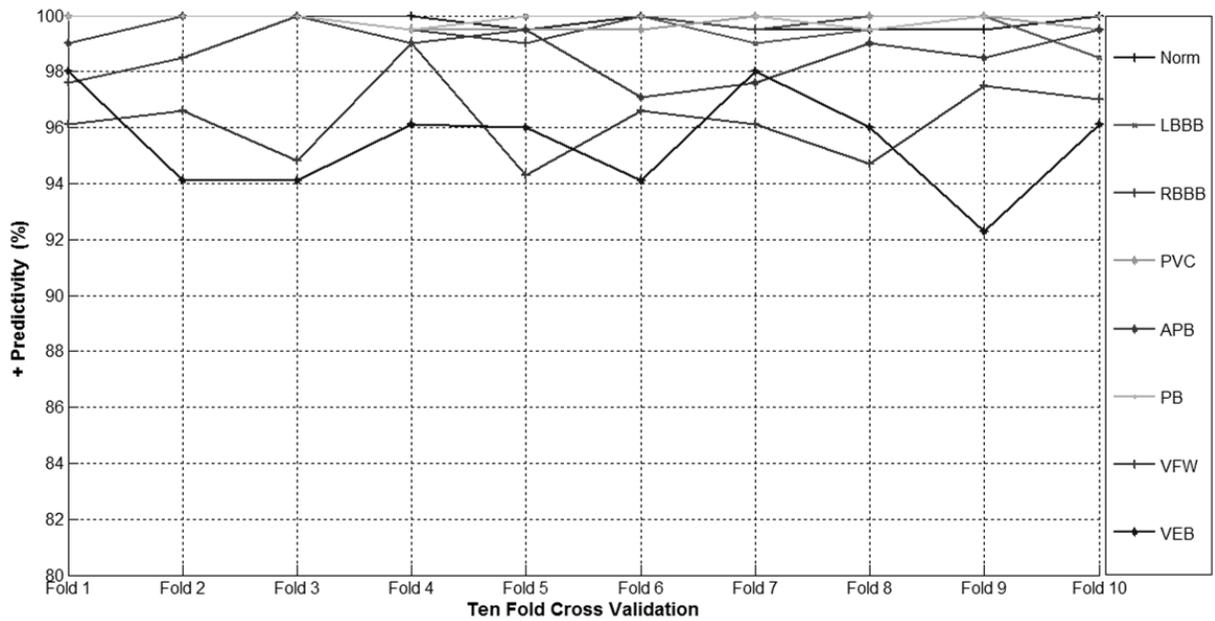


Figure 8-17 Results of positive predictive value with SVM classifier

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
NORM	100.0	100.0	100.0	100.0	99.5	100.0	99.5	99.5	99.5	100.0
LBBB	100.0	100.0	100.0	99.5	100.0	100.0	99.0	99.5	100.0	98.5
RBBB	97.6	98.5	100.0	99.5	99.0	100.0	99.5	100.0	100.0	99.5
PVC	100.0	100.0	100.0	99.5	99.5	99.5	100.0	100.0	100.0	99.5
APB	99.0	100.0	100.0	99.0	99.5	97.1	97.6	99.0	98.5	99.5
PB	100.0	100.0	100.0	99.5	100.0	100.0	100.0	99.5	100.0	100.0
VEB	96.1	96.6	94.8	99.0	94.3	96.6	96.1	94.7	97.5	97.0
VFW	98.0	94.1	94.1	96.1	96.0	94.1	98.0	96.0	92.3	96.1

Table 8-8 Detailed positive predictive value for different arrhythmias with SVM classifier

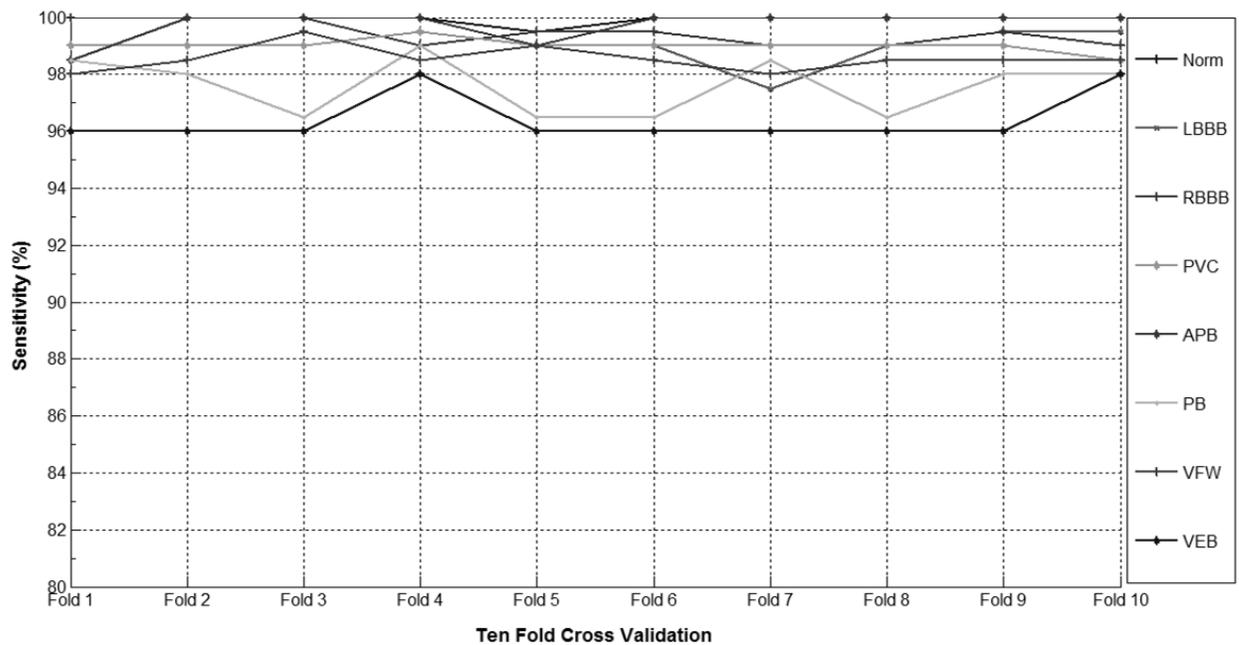


Figure 8-18 Results of sensitivity with SVM classifier

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
NORM	100	100	100	100	99.5	100	100	100	100	100
LBBB	99	99	99	99.5	99	99	97.5	99	99.5	99.5
RBBB	100	100	100	99	99.5	99.5	99	99	99.5	99
PVC	99	99	99	99.5	99	99	99	99	99	98.5
APB	98.5	100	100	100	99	100	100	100	100	100
PB	98.5	98	96.5	99	96.5	96.5	98.5	96.5	98	98
VEB	98	98.5	99.5	98.5	99	98.5	98	98.5	98.5	98.5
VFW	96	96	96	98	96	96	96	96	96	98

Table 8-9 Detailed sensitivity for different arrhythmias with SVM classifier

Figure 8-19 shows the accuracy for different folds of several arrhythmias using principal components features and ANN classifier. It can be noted from this figure, that PCA extracted features set provided less accuracy as compared to proposed features set with independent components as features. Overall accuracy is 98.4 % with VFW beats recorded only 95.6 % . Figure 8-20 shows the positive predictive value of PCA features with ANN classifier for different folds. The sensitivity of classification of PCA components for different folds of the ANN classifier is shown plotted in Figure 8-21. It can be seen from this figure that, NORM beat achieved highest sensitivity consistently for all folds.

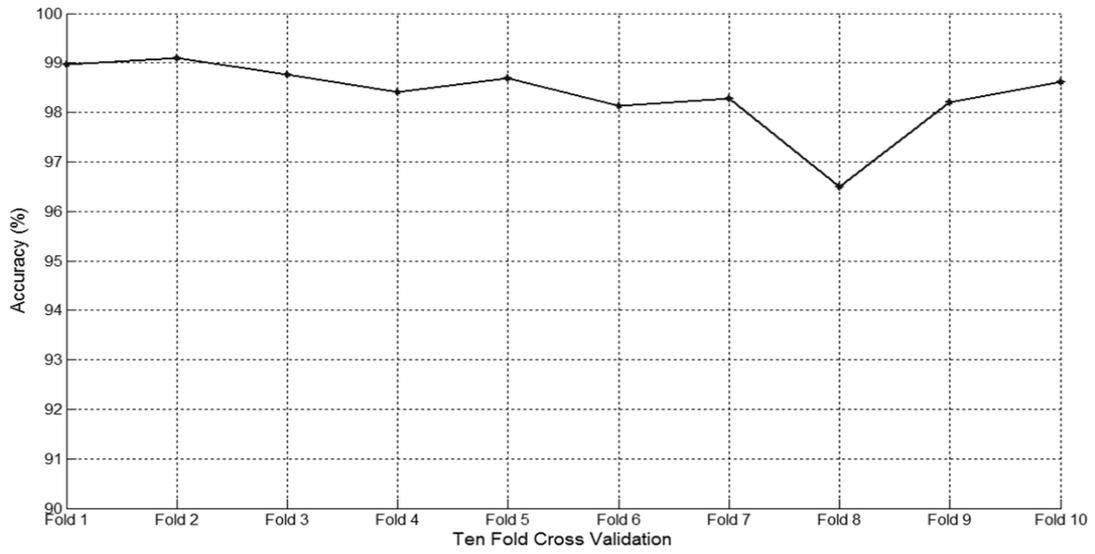


Figure 8-19 Results of accuracy with PCA extracted features

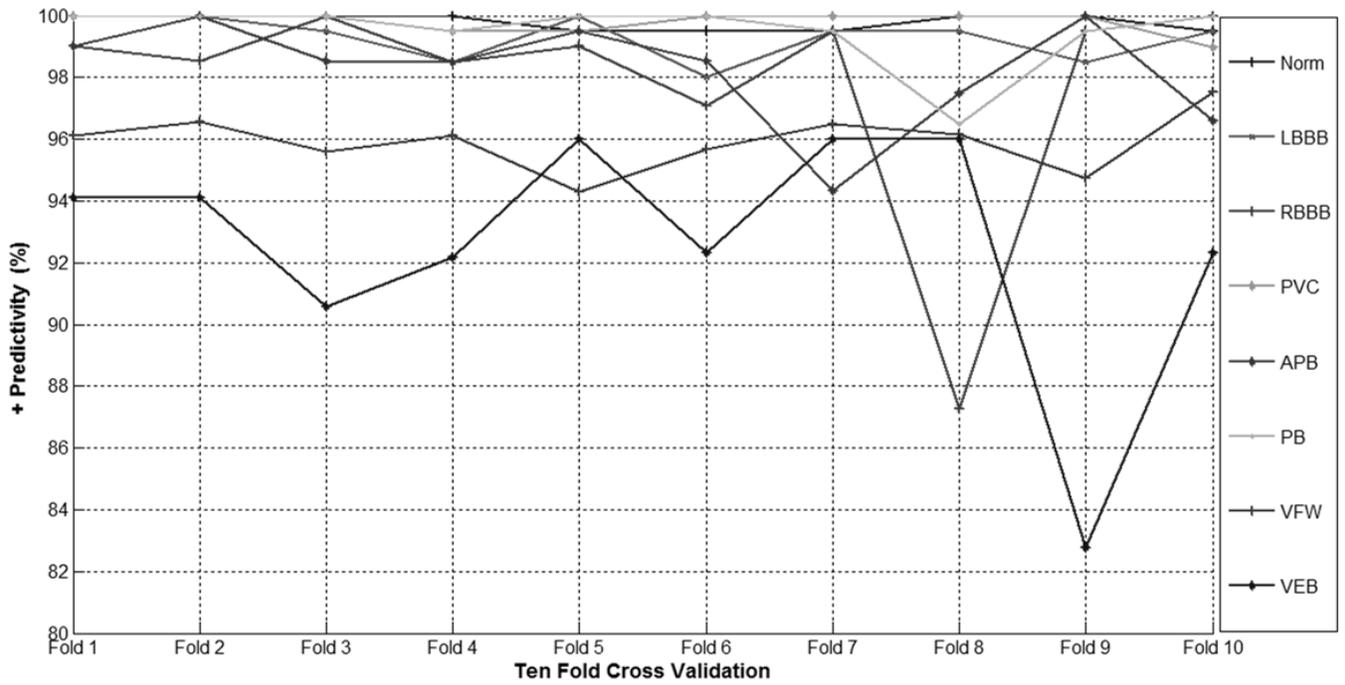


Figure 8-20 Results of positive predictive value with PCA extracted features

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
NORM	100	100	100	100	99.5	100	100	100	100	100
LBBB	99	99	99	99	99	99	93.5	99	99	99
RBBB	100	100	100	99	99.5	99.5	99	99	99	99
PVC	99	99	99	98	99	99	99	99	98.5	98
APB	98.5	100	100	99	99	100	100	100	100	100
PB	98.5	98	96.5	96.5	96.5	96.5	96.5	96.5	96.5	96.5
VEB	98	98.5	99.5	98.5	99	98.5	92.03	98.5	98.5	97.5
VFW	96	96	96	94	96	96	96	96	96	96

Table 8-10 Detailed specificity for different arrhythmias with PCA extracted features

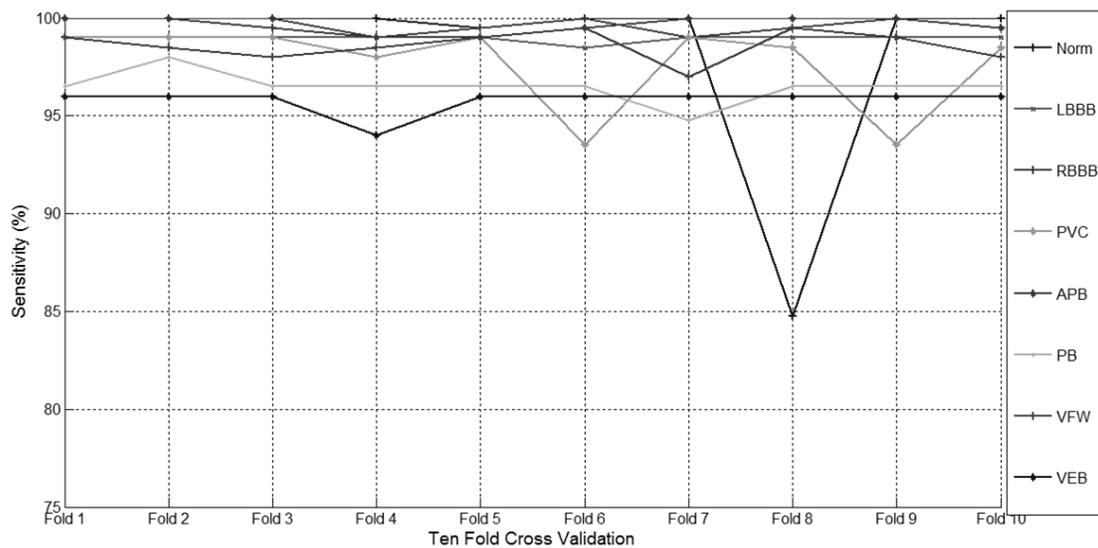


Figure 8-21 Results of positive predictive value with PCA extracted features

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
NORM	100	100	100	100	99.5	100	100	100	100	100
LBBB	99	99	99	99	99	99	93.5	99	99	99
RBBB	100	100	100	99	99.5	99.5	99	99	99	99
PVC	99	99	99	98	99	99	99	99	98.5	98
APB	98.5	100	100	99	99	100	100	100	100	100
PB	98.5	98	96.5	96.5	96.5	96.5	96.5	96.5	96.5	96.5
VEB	98	98.5	99.5	98.5	99	98.5	92.03	98.5	98.5	97.5
VFW	96	96	96	94	96	96	96	96	96	96

Table 8-11 Detailed sensitivity for different arrhythmias with PCA extracted features

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
Accuracy	98.9	99.1	99.0	98.4	98.7	98.8	97.1	98.8	98.7	98.5

Table 8-12 Detailed accuracy for different arrhythmias with PCA extracted features

8.1 Discussion

Efficient artifact reduction in ambulatory ECG recording requires a broad approach. In this work, an integrated approach to motion artifact reduction in ECG combining the electrode position selection and algorithm development using ICA is presented. Initial results in these areas have been reported, and important challenges remain. In order to evaluate the proposed algorithm, ambulatory ECGs were acquired from the MIT-BIH database, and the algorithm is compared with several methods including normally used algorithm.

The selection of MIT-BIH database for this study was due to the fact that most of the previous work mentioned in the literature for ECG classification, were tested only on limited data sets and the generalization performance of these methods on large databases was not tested. Secondly, all these methods were tested only on a few classes of ECG beats and there is a need to test the methods and algorithms on a standard classification scheme of arrhythmia beats such as ANSI/AAMI EC57:1998. Since the proposed algorithm achieved the desired performance, it should be tested on real patients ECG data for real time implementation.

The application of statistical de-noising techniques such as ICA requires independency of the input leads, thus imposing additional and possibly conflicting requirements. Further research is needed to evaluate the most appropriate electrode configuration, as required by cardiologists, optimizing the performances of de-noising algorithms, and improving the comfort of the patients.

For a complete automatic heartbeat classification system, R-peak detection is required as the preliminary step. The errors result from the R-wave detection unit may reduce the performance of the classifier. Since, a number of research work published in the literature achieved impressively high detection rates of more than 99.5% (Afonso, Tompkins, Nguyen, & Luo, 1999; Lagerholm, Peterson, Braccini, Edenbrandt, & Sornmo, 2000). This work does not intend to detect the R-waves, but use the information provided by the annotation files in the MIT-BIH database, which were manually verified by specialists. It is straightforward to incorporate the R-wave detection algorithms into the proposed scheme to make a fully automatic heartbeat classification system.

Results presented above are obtained by processing the ECG in slabs of 10 seconds. Although this is appropriate during the algorithm development and performance optimization course, it does not provide the necessary real-time performances required for continuous and remote monitoring of ECG on-the-move. The easiest way to extend a batch-processing algorithm for real-time applications is to compute it in sliding overlapping windows. But that too has several drawbacks, including reduced algorithmic performance, increased memory requirement and increased computational load.

In summary, the proposed motion artifact removal algorithm using the ICA that minimizes distortion. The proposed method performed better than referential methods in the presence of all common types of artefacts. For ECG classification with features extracted using independent component analysis, the proposed algorithm performs better than any existing method. Also, the algorithm which is based on short-term segmented dataset offers potential for real-time processing this algorithm.

Chapter 9 Conclusion and Future Work

The work described in this thesis has been concerned with the development of a motion artefact denoising algorithm and classification algorithm for ECG. Two new algorithms are developed, one for removing non trivial noises in ECG and second for ECG arrhythmia classification using features extracted with ICA. A new feature set is developed which includes dynamic ECG features along with ICA extracted features.

9.1 Propose and develop the algorithm for denoising of ECG signals using ICA

The proposed algorithm deals with uncommon noises presented in ECG data during holter and telemedicine applications. Proposed method connects well-known FastICA algorithm with a correlation factor and kurtosis in order to identify noise and reduce it. This study investigated the performance of PCA and ICA in denoising ECG signals recorded in ambulatory conditions. A simulated database formed by the combination of clean ECG signals with electrode motion artefacts scaled to different levels of energy is developed for evaluation. Sensitivity of the beat detection algorithm after filtering with ICA is 100 %, even for very noisy signal. High classification accuracy and positive predictive value of ICA filtered signal are recorded. 40 % improvement in the sensitivity of ICA filtered signal for SNR of -12 dB is achieved. An automatic method based on kurtosis and correlation coefficient for component selection is proposed. Filtering, using this method, achieved 100 % sensitivity in beat detection as compared to non-filtered signals, especially when the noise level is high. As a limitation of this study, it should be noted that some stationarity has been assumed, as signals are of 10 seconds length. The performance under

shorter duration noise is not studied. Although this is acceptable during the algorithm development and performance optimization practice, it does not provide the necessary real-time performances required for continuous and remote monitoring of ECG on-the-move. Another limitation to the current study is that it has been tested on a database in which clean ECG and noise signals are artificially combined. Additional tests are needed to be done in real-time situations to understand and expand the scope of the proposed algorithm.

The proposed method performed better than referential methods in the presence of all common types of artefacts. FastICA algorithm has better performance as compared to other ICA algorithms discussed in literature. The time requirements of our algorithm are decreased due computational simplicity of the FastICA algorithm. This ability is very valuable in medical applications.

9.2 Proposed and developed the algorithm for arrhythmia classification using ICA

The algorithm is proposed for ECG arrhythmia classification by using features extracted with ICA. The classification experiments are performed on the MIT-BIH arrhythmia database. For this study, eight types of ECG samples which includes the normal sinus beat and arrhythmias, are used. Proposed method combines Independent Component Analysis (ICA), Pre-RR Interval, post-RR Interval and QRS Segment Power for feature set and neural network classifiers for ECG beat classification. A new feature set combined dynamic ECG features along with ICA extracted features. Optimal selection of the number of ICs for best classification accuracy is made. Computer simulations show that ICA based feature extraction method performs better than any other available method. Neural network classifiers

demonstrated high classification accuracies of over 99.5% with a relatively small number of ICs. It also proves that selecting relevant features from the feature set could improve the recognition performance, which makes a robust feature set for ECG arrhythmia classification. Comparison of classification accuracy, specificity and sensitivity with other state of the art method is done and its efficiency is proven. The results prove that the proposed scheme is a promising model for arrhythmia detection of clinical ECG signals.

All goals formulated were successfully met. Two algorithms were proposed– one for denoising of ECG recording containing trivial noise and the second algorithm for classification of eight different types of ECG arrhythmias. Both algorithms were tested against the state-of-the-art methods and results were decent. Both algorithms were capable of dealing with uncommon noises. This makes them very useful in applications within telemedicine issues and holter recordings, where the environment is rapidly changing and the distortions corrupting the ECG signals could be very different from those normally present in ECG within the laboratory or clinical measurement of ECG.

Future Work

Although the results presented here have demonstrated the effectiveness of the Independent Component Analysis for removal of motion artefacts and feature extraction of ECG signal and it could be further developed in a number of ways:

Extending the Algorithm for Real-Time Application

To extend the proposed batch-processing algorithm for real-time applications, computation should be done in sliding overlapping windows. There are several shortcomings in running algorithms in sliding overlapping windows, including reduced algorithmic performance, increased memory requirement and increased computational load, which needs to be resolved for a streamline implementation.

Extending the Algorithm for Other Classifiers

No classifier is perfect; the best classifier's performance is to correctly classify novel cases. Performance is related to both classifier design and testing. Occasionally complex classifiers fit 'noise' in the training data, achieving low accuracy when presented with novel cases. Our classification results showed a different accuracy of each arrhythmia with same classifier. Using multiple and hybrid classifiers can result in improving classification accuracy and stability. The future work will aim at extensive testing of our developed algorithms with real data sets in conjugation with other intelligent system classifiers.

References

1. Acernese, F., Ciaramella, A., De Martino, S., Falanga, M., Godano, C., & Tagliaferri, R. (2004). Polarisation analysis of the independent components of low frequency events at Stromboli volcano (Eolian Islands, Italy). *Journal of Volcanology and Geothermal Research*, 137(1), 153–168.
2. Acharya, R. (2007). *Advances in cardiac signal processing*. Springer.
3. Acharya, R., Kumar, A., Bhat, P. S., Lim, C. M., Kannathal, N., & Krishnan, S. M. (2004). Classification of cardiac abnormalities using heart rate signals. *Medical and Biological Engineering and Computing*, 42(3), 288–293.
4. Afonso, V. X., Tompkins, W. J., Nguyen, T. Q., & Luo, S. (1999). ECG beat detection using filter banks. *Biomedical Engineering, IEEE Transactions on*, 46(2), 192–202.
5. Al-Fahoum, A. S., & Howitt, I. (1999). Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias. *Medical & Biological Engineering & Computing*, 37(5), 566–573.
6. Anand, S. S., & Yusuf, S. (2011). Stemming the global tsunami of cardiovascular disease. *Lancet*, 377(9765), 529–32. doi:10.1016/S0140-6736(10)62346-X
7. Asl, B. M., Setarehdan, S. K., & Mohebbi, M. (2008a). Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artificial Intelligence in Medicine*, 44(1), 51–64.
8. Asl, B. M., Setarehdan, S. K., & Mohebbi, M. (2008b). Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artificial Intelligence in Medicine*, 44(1), 51–64.

9. Bache, K., & Lichman, M. (2013). *UCI Machine Learning Repository*. University of California, Irvine, School of Information and Computer Sciences. Retrieved from <http://archive.ics.uci.edu/ml>
10. Barro, S., Fernández-Delgado, M., Vila-Sobrino, J. A., Regueiro, C. V., & Sánchez, E. (1998). Classifying multichannel ECG patterns with an adaptive neural network. *IEEE Engineering in Medicine and Biology Magazine: The Quarterly Magazine of the Engineering in Medicine & Biology Society*, 17(1), 45–55.
11. Barros, A. K., Mansour, A., & Ohnishi, N. (1998). Removing artefacts from electrocardiographic signals using independent components analysis. *Neurocomputing*, 22(1), 173–186.
12. Belgacem, N., Chikh, M. A., & Reguig, F. B. (2003a). Supervised classification of ECG using neural networks. Retrieved from <http://dspace.univ-tlemcen.dz/handle/112/837>
13. Belgacem, N., Chikh, M. A., & Reguig, F. B. (2003b). Supervised classification of ECG using neural networks. Retrieved from <http://dspace.univ-tlemcen.dz/handle/112/837>
14. Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6), 1129–1159.
15. Bingham, E., Kuusisto, J., & Lagus, K. (2002). ICA and SOM in text document analysis. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 361–362). ACM.

16. Cabras, G., Carniel, R., & Isserman, J. (2010). Signal enhancement with generalized ICA applied to Mt. Etna volcano, Italy. *Bollettino Di Geofisica Teorica Ed Applicata*, 51(1).
17. Cardoso, J.-F., & Souloumiac, A. (1993). Blind beamforming for non-Gaussian signals. In *IEE Proceedings F (Radar and Signal Processing)* (Vol. 140, pp. 362–370). IET. Retrieved from <http://digital-library.theiet.org/content/journals/10.1049/ip-f-2.1993.0054>
18. Cardoso, J.-F., & Souloumiac, A. (1996). Jacobi angles for simultaneous diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 17(1), 161–164.
19. Castells, F., Laguna, P., Sö, L., Bollmann, A., & Roig, M. (2007). Principal component analysis in ECG signal processing. *EURASIP Journal on Advances in Signal Processing*, 2007.
20. Castells, F., Rieta, J. J., Millet, J., & Zarzoso, V. (2005). Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias. *Biomedical Engineering, IEEE Transactions on*, 52(2), 258–267.
21. Chawla, M. P. S. (2009). A comparative analysis of principal component and independent component techniques for electrocardiograms. *Neural Computing and Applications*, 18(6), 539–556.
22. Cheung, Y., & Xu, L. (2001). Independent component ordering in ICA time series analysis. *Neurocomputing*, 41(1), 145–152.
23. Chou, K.-T., & Yu, S.-N. (2008). Categorizing Heartbeats by Independent Component Analysis and Support Vector Machines. In *Intelligent Systems Design and*

- Applications, 2008. ISDA '08. Eighth International Conference on* (Vol. 1, pp. 599–602).
24. Cichocki, A., Amari, S., Siwek, K., Tanaka, T., & Phan, A. H. (2003). ICALAB for Signal Processing. *Toolbox for ICA, BSS and BSE*, [http://www. Bsp. Brain. Riken. ip/\[ICALAB/ICALAB SignalProc](http://www.Bsp.Brain.Riken.ip/ICALAB/ICALAB SignalProc).
25. Clifford, G. D., Azuaje, F., & McSharry, P. (2006). *Advanced methods and tools for ECG data analysis*. Artech house London.
26. Davie, A. P., Francis, C. M., Love, M. P., Caruana, L., Starkey, I. R., Shaw, T. R. D., ... McMurray, J. J. V. (1996). Value of the electrocardiogram in identifying heart failure due to left ventricular systolic dysfunction. *Bmj*, *312*(7025), 222.
27. De Chazal, P., O'Dwyer, M., & Reilly, R. B. (2004). Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *Biomedical Engineering, IEEE Transactions on*, *51*(7), 1196–1206.
28. De, O., Adams, M., Carnethon, G., Lloyd-Jones, R., Simone, T., Ferguson, K., ... Hong American. (2009). D. Association Statistics Committee and Stroke Statistics Subcommittee, heart disease and stroke statistics- update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee, *Circulation* *119* (January (3)) ., 480–486.
29. Dokur, Z., & Ölmez, T. (2001). ECG beat classification by a novel hybrid neural network. *Computer Methods and Programs in Biomedicine*, *66*(2), 167–181.

30. EC57, A.-A. (1998). Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. *Association for the Advancement of Medical Instrumentation, Arlington, VA.*
31. Fikret, G., & EMB. (1999). Neural network based decision making in diagnostic applications”,*IEEE* , , pp. ., *18*(4), 89–93.
32. Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... Stanley, H. E. (2000). Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. *Circulation*, *101*(23), e215–e220.
33. Güler, İ. (2005). ECG beat classifier designed by combined neural network model. *Pattern Recognition*, *38*(2), 199–208.
34. Güler, İ., & Übeyli, E. D. (2005). A modified mixture of experts network structure for ECG beats classification with diverse features. *Engineering Applications of Artificial Intelligence*, *18*(7), 845–856.
35. Guvenir, H. A., Acar, S., Demiroz, G., & Cekin, A. (1997). A supervised machine learning algorithm for arrhythmia analysis. In *Computers in Cardiology 1997* (pp. 433–436). doi:10.1109/CIC.1997.647926
36. Haykin, S. (1995). *Neural Networks: Foundation*. MacMillan College Publishing Company,.
37. Holter, N. J. (1961). New Method for Heart Studies Continuous electrocardiography of active subjects over long periods is now practical. *Science*, *134*(3486), 1214–1220.

38. Homaeinezhad, M. R., Atyabi, S. A., Tavakkoli, E., Toosi, H. N., Ghaffari, A., & Ebrahimpour, R. (2012). ECG arrhythmia recognition via a neuro-SVM–KNN hybrid classifier with virtual QRS image-based geometrical features. *Expert Systems with Applications*, *39*(2), 2047–2058.
39. Hu, W., Xie, D., & Tan, T. (2004). A hierarchical self-organizing approach for learning the patterns of motion trajectories. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council*, *15*(1), 135–44.
doi:10.1109/TNN.2003.820668
40. Huang, H. F., Hu, G. S., & Zhu, L. (2012). Sparse representation-based heartbeat classification using independent component analysis. *Journal of Medical Systems*, *36*(3), 1235–1247.
41. Huikuri, H. V., Castellanos, A., & Myerburg, R. J. (2001). Sudden death due to cardiac arrhythmias. *The New England Journal of Medicine*, *345*(20), 1473–82.
doi:10.1056/NEJMra000650
42. Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *Neural Networks, IEEE Transactions on*, *10*(3), 626–634.
43. Hyvärinen, A. (1999). The fixed-point algorithm and maximum likelihood estimation for independent component analysis. *Neural Processing Letters*, *10*(1), 1–5.
44. Hyvärinen, A., Karhunen, J., & Oja, E. (2001). What is Independent Component Analysis? *Independent Component Analysis*, 145–164.
45. Hyvärinen, A., & Oja, E. (2000a). Independent component analysis: algorithms and applications. *Neural Networks*, *13*(4), 411–430.

46. Hyvärinen, A., & Oja, E. (2000b). Independent component analysis: algorithms and applications. *Neural Networks*, 13(4), 411–430.
47. Jang, J.-S. R., Sun, C.-T., & Mizutani, E. (1997a). Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]. *Automatic Control, IEEE Transactions on*, 42(10), 1482–1484.
48. Jang, J.-S. R., Sun, C.-T., & Mizutani, E. (1997b). Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]. *Automatic Control, IEEE Transactions on*, 42(10), 1482–1484.
49. Jiang, X., Zhang, L., Zhao, Q., & Albayrak, S. (2006a). ECG arrhythmias recognition system based on independent component analysis feature extraction. In *TENCON 2006. 2006 IEEE Region 10 Conference* (pp. 1–4).
50. Jiang, X., Zhang, L., Zhao, Q., & Albayrak, S. (2006b). ECG arrhythmias recognition system based on independent component analysis feature extraction (pp. 1–4).
51. Jung, T.-P., Humphries, C., Lee, T.-W., Makeig, S., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (1998). Removing electroencephalographic artefacts: comparison between ICA and PCA. In *Neural Networks for Signal Processing VIII, 1998. Proceedings of the 1998 IEEE Signal Processing Society Workshop* (pp. 63–72). IEEE.
52. Jung, T.-P., Makeig, S., Lee, T.-W., McKeown, M. J., Brown, G., Bell, A. J., & Sejnowski, T. J. (2000). Independent component analysis of biomedical signals (pp. 633–644).

53. Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artefacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, *111*(10), 1745–1758.
54. Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2001). Analysis and visualization of single-trial event-related potentials. *Human Brain Mapping*, *14*(3), 166–185.
55. Karhunen, J., & Oja, E. (2001). Hyvaˆrinen, A., & Independent component analysis. *John Wiley Sons*.
56. Keerthi, S. S., & Lin, C.-J. (2003). Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Computation*, *15*(7), 1667–1689.
57. Khawaja, A. (2006). *Automatic ECG analysis using principal component analysis and wavelet transformation*. Univ.-Verlag Karlsruhe. Retrieved from <http://www.uvka.de/univerlag/volltexte/2007/227/>
58. Kohonen, T. (1987). Adaptive, associative, and self-organizing functions in neural computing., *26*(23), 4910–8.
59. Kohonen, T. (1988). Self-organization and associative memory. *Self-Organization and Associative Memory*, *100 Figs. XV, 312 Pages.. Springer-Verlag Berlin Heidelberg New York. Also Springer Series in Information Sciences, Volume 8, 1.*
60. Kolenda, T., Hansen, L. K., & Sigurdsson, S. (2000). Independent components in text. In *Advances in Independent Component Analysis* (pp. 235–256). Springer.

61. Kwak, N., & Choi, C.-H. (2003). Feature extraction based on ICA for binary classification problems. *Knowledge and Data Engineering, IEEE Transactions on*, 15(6), 1374–1388.
62. Kwak, N., Choi, C.-H., & Choi, J. Y. (2001). Feature extraction using ica. In *Artificial Neural Networks—ICANN 2001* (pp. 568–573). Springer. Retrieved from http://link.springer.com/chapter/10.1007/3-540-44668-0_80
63. Li, Y., Powers, D., & Peach, J. (2000). Comparison of blind source separation algorithms. *Advances in Neural Networks and Applications*, 18–21.
64. Linh, S., Osowski, M., & IEEE. (2003). T.H. Stodolski, On-line heart beat recognition using Hermite polynomials and neuro-fuzzy network, on *Instrumentation and Measurement* 52 (August (4)) ., 1224–1231.
65. Lu, C.-J., Lee, T.-S., & Chiu, C.-C. (2009). Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems*, 47(2), 115–125.
66. Madhuranath, H., & Haykin, S. (1998). *Improved Activation Functions for Blind Separation: Details of Algebraic Derivations*. CRL Internal Report No.
67. Makeig, S., Bell, A. J., Jung, T.-P., Sejnowski, T. J., & others. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems*, 145–151.
68. Makeig, S., Jung, T.-P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences*, 94(20), 10979–10984.

69. Malmivuo, J., & Plonsey, R. (1995). *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press.
70. Martis, R. J., Acharya, U. R., & Min, L. C. (2013). ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform. *Biomedical Signal Processing and Control*, 8(5), 437–448.
71. Matei, B. (2000). A review of independent component analysis techniques. *Signal*, 6(7), 8.
72. McGarry, K., Sarfraz, M., & MacIntyre, J. (2007). Integrating gene expression data from microarrays using the self-organising map and the gene ontology. In *Pattern Recognition in Bioinformatics* (pp. 206–217). Springer.
73. Meyer, C. D. (2000). *Matrix analysis and applied linear algebra*. Siam.
74. Minami, K., Nakajima, H., & Toyoshima, T. (1999). Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network. *Biomedical Engineering, IEEE Transactions on*, 46(2), 179–185.
75. Mishra, P., & Singla, S. K. (2013). Artefact Removal from Biosignal using Fixed Point ICA Algorithm for Pre-processing in Biometric Recognition. *Measurement Science Review*, 13(1), 7–11.
76. Mohammadzadeh-Asl, B., & Setarehdan, S. K. (2006). Neural network based arrhythmia classification using heart rate variability signal. In *Proceedings of the EUSIPCO*.

77. Moody, G. B., Mark, R. G., & Goldberger, A. L. (2001). PhysioNet: a web-based resource for the study of physiologic signals. *Engineering in Medicine and Biology Magazine, IEEE*, 20(3), 70–75.
78. Moody, G. B., Muldrow, W., & Mark, R. G. (1984). A noise stress test for arrhythmia detectors. *Computers in Cardiology*, 11(3), 381–384.
79. Naik, G. R., & Kumar, D. K. (2011). An overview of independent component analysis and its applications. *Informatica: An International Journal of Computing and Informatics*, 35(1), 63–81.
80. Naik, G. R., Kumar, D. K., & Arjunan, S. P. (2010). Independent Component Analysis For Classification Of Surface Electromyography Signals During Different MVCs. In *Proceedings of the 20th International EURASIP Conference-BIOSIGNAL 2010: Analysis of Biomedical Signals and Images* (pp. 352–358).
81. Nazmy, T. M., El-Messiry, H., & Al-Bokhity, B. (2010). Adaptive neuro-fuzzy inference system for classification of ECG signals (pp. 1–6). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5461743
82. Negnevitsky, M. (2005). *Artificial intelligence: a guide to intelligent systems*. Pearson Education.
83. Olmez, T. (1997). Classification of ECG waveforms using RCE neural network and genetic algorithm', *Electronics Letters*, , pp. ., 33(8), 1561–1562.
84. Owis, M. I., Youssef, A.-B., & Kadah, Y. M. (2002). Characterisation of electrocardiogram signals based on blind source separation. *Medical and Biological Engineering and Computing*, 40(5), 557–564.

85. Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *Biomedical Engineering, IEEE Transactions on*, (3), 230–236.
86. Pu, Q., & Yang, G.-W. (2006). Short-text classification based on ICA and LSA. In *Advances in Neural Networks-ISNN 2006* (pp. 265–270). Springer. Retrieved from http://link.springer.com/chapter/10.1007/11760023_39
87. Qiu, G., Varley, M. R., & Terrell, T. J. (1992). Accelerated training of backpropagation networks by using adaptive momentum step. *Electronics Letters*, 28(4), 377–379.
88. Rice, J. A. (2007). *Mathematical statistics and data analysis*. Cengage Learning.
89. Romero, I. (2010). PCA-based noise reduction in ambulatory ECGs (pp. 677–680).
90. Romero, I. (2011). PCA and ICA applied to Noise Reduction in Multi-lead ECG (pp. 613–616).
91. Rosaria, S., Carlo, M., & IEEE. (1998). Artificial neural networks for automatic ECG analysis', *Signal Processing*, , pp. ., 46(5), 1417–1425.
92. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). *Learning Internal Representations by Error Propagation, Parallel Distributed Processing, Explorations in the Microstructure of Cognition, ed. DE Rumelhart and J. McClelland. Vol. 1. 1986*. MIT Press, Cambridge, MA.
93. Sarfraz, M., & Li, F. (2013). Independent Component Analysis for Motion Artefacts Removal from Electrocardiogram. *Global Perspectives on Artificial Intelligence*, 1(4).
94. Sarfraz, M., Li, F., & Javed, M. (2011). A comparative study of ICA algorithms for ECG signal processing (pp. 135–138).

95. Sarfraz, M., Li, F., & Javed, M. (2013). A Blind Source Separation Method to Eliminate Noise Artefacts in ECG Signals (Vol. I, pp. 112–119). Presented at the 2nd INTERNATIONAL CONFERENCE ON RECENT TRENDS IN COMPUTING, Ghaziabad, India.
96. Sarfraz, M., Li, F., & Khan, A. A. (2014). Independent Component Analysis Methods to Improve Electrocardiogram Patterns Recognition in the Presence of non-Trivial Artefacts. In *ICBBS 2014* (Vol. III). Copenhagen, Denmark.
97. Sheffield, L. T., Berson, A., Bragg-Remschel, D., Gillette, P. C., Hermes, R. E., Hinkle, L., ... Oliver, C. (1985). Recommendations for standards of instrumentation and practice in the use of ambulatory electrocardiography. The Task Force of the Committee on Electrocardiography and Cardiac Electrophysiology of the Council on Clinical Cardiology. *Circulation*, *71*(3), 626A.
98. Shen, M., Wang, L., Zhu, K., & Zhu, J. (2010). Multi-lead ECG classification based on independent component analysis and support vector machine. In *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on* (Vol. 3, pp. 960–964). IEEE.
99. Shlens, J. (2014). A Tutorial on Independent Component Analysis. *arXiv Preprint arXiv:1404.2986*.
100. Smith, L. I. (2002). A tutorial on principal components analysis. *Cornell University, USA*, *51*, 52.
101. Stone, J. V. (2004). *Independent component analysis: a tutorial introduction*. The MIT Press.

102. Tanskanen, J., Mikkonen, J. E., & Penttonen, M. (2005). Independent component analysis of neural populations from multielectrode field potential measurements. *Journal of Neuroscience Methods*, *145*(1), 213–232.
103. Tichavský, P., Zima, M., & Krajca, V. (2011). Automatic Removal of Sparse Artefacts in Electroencephalogram. In *BIOSIGNALS* (pp. 530–535). Vapnik, V., Golowich, S. E., & Smola, A. (1997). Support vector method for function approximation, regression estimation, and signal processing. *Advances in Neural Information Processing Systems*, 281–287.
104. Wang, J.-S., Chiang, W.-C., Yang, Y.-T. C., & Hsu, Y.-L. (2012). An effective ECG arrhythmia classification algorithm. In *Bio-Inspired Computing and Applications* (pp. 545–550). Springer.
105. Wisbeck, J. O., Barros, A. K., Yy, A. K. B., & Ojeda, R. G. (1998). Application of ICA in the separation of breathing artefacts in ECG signals. Wu, S., Chow, T., Using, S., RBF, & IEEE. (2004). Induction Machine Fault Detection Based Networks', *Industrial Electronics*, pp. ., *51*, 183–194.
106. Wu, Y., & Zhang, L. (2011). ECG classification using ICA features and support vector machines. In *Neural Information Processing* (pp. 146–154). Springer.
107. Yu, C.-C., & Liu, B.-D. (2002). A backpropagation algorithm with adaptive learning rate and momentum coefficient. In *Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on* (Vol. 2, pp. 1218–1223). IEEE.

108. Yu, S.-N., & Chou, K.-T. (2006). Combining independent component analysis and backpropagation neural network for ECG beat classification (pp. 3090–3093).
109. Yu, S.-N., & Chou, K.-T. (2007). A switchable scheme for ECG beat classification based on independent component analysis. *Expert Systems with Applications*, 33(4), 824–829.
110. Yu, S.-N., & Chou, K.-T. (2008). Integration of independent component analysis and neural networks for ECG beat classification. *Expert Systems with Applications*, 34(4), 2841–2846.
111. Yu, S.-N., & Chou, K.-T. (2009). Selection of significant independent components for ECG beat classification. *Expert Systems with Applications*, 36(2), 2088–2096.
112. Zhu, Y., Shayan, A., Zhang, W., Chen, T. L., Jung, T.-P., Duann, J.-R., ... Cheng, C.-K. (2008). Analyzing high-density ECG signals using ICA. *Biomedical Engineering, IEEE Transactions on*, 55(11), 2528–2537.

Appendix: Related Publication

- Sarfraz, M., Li, F., & Javed, M. (2011, July). A comparative study of ICA algorithms for ECG signal processing. In Proceedings of the International Conference on Advances in Computing and Artificial Intelligence (pp. 135-138). ACM.
- Sarfraz, M., & Li, F. (2013). Independent Component Analysis for Motion Artefacts Removal from Electrocardiogram. Global Perspectives on Artificial Intelligence, Vol. 1 Iss.4, (pp. 49-55) .
- Sarfraz, M., Li, F., & Javed, M. (2013). A Blind Source Separation Method to Eliminate Noise Artefacts in ECG Signals (Vol. I, pp. 112–119). Presented at the 2nd international conference on recent trends in computing, Ghaziabad, India.
- Sarfraz, M., Li, F., & Khan, A. A. (2014). Independent Component Analysis Methods to Improve Electrocardiogram Patterns Recognition in the Presence of non-Trivial Artefacts. International Journal of Bioscience, Biochemistry and Bioinformatics (IJBBB ISSN: 2010-3638) (*In Press*).
- Sarfraz, M., Li, F., & Khan, A. A. (2014). Using Independent Component Analysis to Obtain Feature Space for Reliable ECG Arrhythmia Classification. Manuscript submitted for publication.