FUSION OF ECG/EEG FOR IMPROVED AUTOMATIC SEIZURE DETECTION USING DEMPSTER SHAFER THEORY OF EVIDENCE

BY

MOHAMMED ABDUL AZEEM SIDDIQUI

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

ELECTRICAL ENGINEERING

May 2011

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **MOHAMMED ABDUL AZEEM SIDDIQUI** under the direction of his thesis advisor and approved by his thesis committee, has been presented to and accepted by Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN ELECTRICAL ENGINEERING**.

Thesis Committee

Dr. Mohamed A. Deriche (Advisor)

Dr. Mohamed Mohandes (Co-Advisor)

Dr. Abdelmalek Zidouri(Member)

Dr. Nabeel Maalej (Member)

Dr. Samer M.Arafat (Member)

Dr. Ali Ahmad Al-Shaikhi
(Department Chairman)

Dr. Salam Adel Zummo
(Dean of Graduate Studies)

عمادة الدراسات العليا
DEANSHIP OF GRADUATE STUDIES

Date 23/8/11

Dedicated to

My Parents

Mr. & Mrs. Mohammed Abdul Gafoor Siddiqui

Whose Prayers and Perseverance led to this accomplishment

ACKNOWLEDGEMENTS

"Read! In the Name of your Lord who created. He has created man from a clot. Read! And your Lord is the Most Generous. Who has taught by the pen. He has taught man that which he knew not." [Al Quran 96 Ayah 1-5]

In the name of Allah, the most gracious and the most merciful. All praise is due to Allah; we praise him; we worship him alone without associating any partners and seek forgiveness from him. Peace and blessings be upon his last messenger Muhammad (saws), his family, his companions, and all those who followed him until Day of judgment.

First and foremost gratitude is due to the esteemed university, the King Fahd University of Petroleum & Minerals for my admittance, and to its learned faculty members for imparting quality learning and knowledge with their valuable support and able guidance that has led my way through this point of undertaking my research work.

My deep appreciation and heartfelt gratitude goes to my thesis advisor Dr. Mohamed A. Deriche for his constant support, encouragement and guidance throughout my thesis work. I would also like to thank my Co-Advisor Dr. Mohamed Mohandes along with the other committee members Dr. Abdelmalek Zidouri, Dr. Nabil Maalej and Dr. Sameer Arafat for their extraordinary and thought provoking contribution in my research. It was surely an honor and exceptional learning to work with all of them.

I owe thanks to my friends, colleagues who made my work and stay at KFUPM very pleasant and joyful. A few of them are Fasi bhaijan ,Saad bhaijan, Ajmal, Abdul malik bhai, Amer, Naeem, Rizwan , Irfan, Akber, Mumtaz bhai, Abdur rahman bhai, Touseef, Afzal, Misbah bhai, Javed, Salman, Najam, Zameer, Mohsin, Wajahat, Sameer, Khaleel tamil, Khaleel and many others of whom I will not be able to name here.

I would like to thank my parents and other family members including all my uncles, aunts and my cousins from the core of my heart. Their prayers and encouragement always help me take the right steps in my life.

May Allah help us in following Islam according to Quran and Sunnah as understood by the Ahlus Sunnah Wal Jamah in the first three generations of Muslim Ummah (Aameen)

Table of Contents

ACKNOWLEDGEMENTS i

LIST OF FIGURES viii

NOMENCLATURE xi

Abbreviations xi

THESIS ABSTRACT xiii

THESIS ABSTRACT (ARABIC) xv

CHAPTER 1 1

INTRODUCTION 1

1.1 Introduction 1

1.2 Some Basic Definitions 2

1.3 Causes of Seizures 5

1.4 Different types of seizures 7

1.5 Dangers of Seizures 8

1.6 Problem Statement 9

1.7 Research Objectives 10

1.8 Organization of Thesis 10

1.9 Section Summary 11

CHAPTER 2 12

LITERATURE REVIEW 12

2.1 Introduction 12

2.2 Biomedical Signal Processing 12

2.3 Seizure detection based on Electroencephalogram (EEG) 13

2.4 Seizure detection based on Electrocardiogram (ECG) 18

2.5 Seizure Detection Based on Other Methods 20

2.6 Combination of Seizure Detection Algorithm 21

2.7 Section Summary 25

CHAPTER 3 26

SEIZURE DETECTION BASED ON EEG SIGNAL 26

3.1 Introduction 26

3.2 EEG Data 27

3.3 Type and Nature of EEG trace 29

3.4 Time Frequency Representation (TFR) 30

3.4.1 Short Time Fourier Transform (STFT) 30

3.4.2 Wigner Ville Distribution (WVD) 33

3.4.3 Choi Williams Distribution 36

3.4.4 Zhao Atlas Marks Distribution (ZAM) 39

3.4.5 Comparison and Conclusion 41

3.5 Singular Value Decomposition 44

3.6 Extracting Feature Vector 45

3.6.1 Left Singular Vectors as Feature Vectors 46

3.6.2 Algorithm for Seizure Detection 48

3.7 Classification 53

3.7.1 Linear Discriminant Analysis 53

3.8 Experimental Results and Performance Comparision 56

3.9 SECTION SUMMARY 58

CHAPTER 4 60

SEIZURE DETECTION BASED ON ECG SIGNAL 60

4.1 Introduction 60

4.2 Anatomy of the Heart 60

4.3 Measurement of Electrical Activity Using ECG 62

4.4 Effects of Seizures on ECG Pattern 65

4.5 ECG database 66

4.6 Extraction of Features from ECG Signals 67

4.6.1 Wavelet Decomposition of ECG Signal: 67

4.6.2 Feature Extraction Algorithm: 72

4.7 Flow Chart of Seizure Detection Algorithm 75

4.8 Classification using Linear Discrimination Analysis 76

4.9 RESULTS AND COMPARISION 77

4.10 SECTION SUMMARY 79

CHAPTER 5 81

COMBINATION OF EEG/ECG USING DEMPSTER SHAFER THEORY OF EVIDENCE .. 81

5.1 Introduction 81

5.2 Different approaches for combination of classifiers 81

5.2.1 Combination of features (Early integration of classifiers (EI)) 82

5.2.2 Combination of classifiers (Late integration of classifiers (LI)) 82

5.3 Types of Combination of Classifiers 83

5.4 Abstract level Combination 84

5.4.1 Majority voting 84

5.4.2 Bagging and Boosting 86

5.4.3 Behavior Knowledge Space 86

5.4.4 Bayesian Formulation 87

5.4.5 Dempster Shafer formulation 87

5.5 Rank level Combination 88

5.6 Measurement level Combination 88

5.6.1 Stacked generalization method 89

5.6.2 Statistical combination method 89

5.6.3 Dempster Shafer theory of combination 89

5.7 Problem of Uncertainty 90

5.8 Dempster Shafer Theory of Evidence 92

5.8.1 Basic belief assignment (BBA) 92

5.8.2 Belief function 93

5.8.3 Plausibility 93

5.8.4 Combination rule 94

5.9 Example 94

5.10 Dempster Shafer combination Algorithm 97

5.11 Combined classification result 101

5.12 Degree of Association 104

5.13 Summary 105

CHAPTER 6 107

FUTURE WORK AND CONCLUSIONS 107

6.1 Future Work 108

References 110

Curriculum Vitae 122

LIST OF FIGURES

CHAPTER 1

Figure 1. 1: Lateral view of Brain [8] 2

Figure 1. 2: A Boy undergoing tonic-clonic seizure [12] 4

CHAPTER 2

Figure 2. 1: Early fusion of features 22

Figure 2. 2: Late fusion of features 22

Figure 2.3: Fusion of probabilities 23

Figure 2.4: Fusion of decisions 23

CHAPTER 3

Figure 3. 1: Standard 10-20 electrode for recording [**46**] 27

Figure 3. 2: Sample EEG signals for non seizure (top) and seizure traces (bottom) 28

Figure 3. 3: STFT of seizure trace with a window size of 150 bins 31

Figure 3. 4:STFT of EEG seizure trace with a window of size 300 bins 32

Figure 3. 5: STFT of EEG seizure trace with a window of size 500 bins 32

Figure 3. 6:Wigner Ville TFR for EEG seizure trace with a window of size 150 bins 34

Figure 3.7:Wigner Ville TFR for EEG seizure trace with a window of size 300 bins 35

Figure 3.8:Wigner Ville TFR for EEG seizure trace with a window of size 500 bins 35

Figure 3.9: Choi Williams TFR for EEG seizure trace with a window of size 150 bins 38

Figure 3.10:Choi Williams TFR for EEG seizure trace with a window of size 300 bins 38

Figure 3.11:Choi Williams TFR for EEG seizure trace with a window of size 500 bins 39

Figure 3. 12: ZAM TFR for EEG seizure trace with a window of size 150 bins 40

Figure 3. 13: ZAM TFR for EEG seizure trace with a window of size 300 bins 41

Figure 3.14: ZAM TFR for EEG seizure trace with a window of size 500 bins 41

Figure 3. 15: STFT TFR for EEG non seizure trace (left) and seizure trace (right) 42

Figure 3. 16: Wigner Ville TFR for EEG non seizure trace (left) and seizure trace (right) 42

Figure 3. 17: Choi Williams TFR for EEG non seizure trace (left) and seizure trace (right) 43

Figure 3. 18: ZAM TFR for EEG non seizure trace (left) and seizure trace (right) 43

Figure 3. 19: Energy of the Singular values of TFR...45

Figure 3.20:Histogram binss of of EEG trace for seizure and its time shifted version..47

Figure 3. 21: (Sample 1) Pmf's of Left and Right singular vector corresponding to 1st singular value of a seizure (Left) and non seizure trace (Right)..50

Figure 3. 22: (Sample 1) Pmf's of Left and Right singular vector corresponding to 1st singular value of a seizure (Left) and non seizure trace (Right)..50

Figure 3. 23: (Sample 2) Pmf's of Left and Right singular vector corresponding to 2nd singular value of a seizure (Left) and non seizure trace (Right)..51

Figure 3.24: (Sample 2) Pmf's of Left and Right singular vector corresponding to 2nd singular value of a seizure (Left) and non seizure trace (Right)..51

Figure 3. 25: Flow chart for feature extraction from EEG signal..52

Figure 3. 26:Representation of Class separation in LDA 54

Figure 3. 27: Seizure detection accuracy as a function of the number of features from LDA....................................56

CHAPTER 4

Figure 4. 1: Heart Valves [**60**]..61

Figure 4. 2: Heart Valves [**60**]..62

Figure 4. 3: ECG waveform [**64**]..64

Figure 4.4: Original ECG signal...66

Figure 4. 5: Wavelet Decomposition tree for ECG signal...69

Figure 4. 6: Types of Biorthogonal wavelets in MATLAB [**75**] 70

Figure 4. 7 Wavelet transformed ECG signal at different levels...71

Figure 4. 8: Filtered and Baseline wander corrected ECG signal..72

Figure 4. 9: Different steps in filtering ECG signal...73

Figure 4. 10 Detected PQRST peaks from the ECG signal 74

Figure 4. 11: Flow chart for ECG feature extraction...75

Figure 4. 12: Seizure detection accuracy as a function of the number of features from LDA....................................78

CHAPTER 5

Figure 5. 1: Combination of features (Early Intergration) 82

Figure 5. 2: Combination of Classifiers (Late integration) 83

Figure 5. 3: Flow Chart for Combining results of ECG/EEG using Dempster Shafer theory of Evidence 100

Figure 5. 4: Receiver Operating Characteristics (ROC) for Case 1 103

Figure 5. 5: : Receiver Operating Characteristics (ROC) for Case 2 103

NOMENCLATURE

Abbreviations

AV	Atrioventricular node
BBA	Basic Belief Assignment
Bel	Belief
BKS	Behavior Knowledge Space
DST	Dempster Shafer Theory
EEG	Electroencephalogram
EI	Early Integration of classifiers
ECG	Electrocardiogram
LDA	Linear Discriminant Analysis
LI	Late Integration of classifiers
PCA	Principal Component Analysis
Pl	Plausibility
STFT	Short Time Fourier Transform
SUDEP	Sudden Unexpected Death in Epilepsy

SA	Sinuatrial node
SVD	Singular Value Decompostion
TF	Time Frequency
WT	Wavelet Transform
ZAM	Zhao Atlas Marks Distribution

THESIS ABSTRACT

Name: Mohammed Abdul Azeem Siddiqui

Title: FUSION OF ECG/EEG FOR IMPROVED AUTOMATIC SEIZURE DETECTION USING DEMPSTER SHAFER THEORY OF EVIDENCE

Major Field: ELECTRICAL ENGINEERING

Date of Degree: May 2011

Objective:

A Dempster Shafer based combination method is presented for the seizure detection algorithm using Electroencephalogram (EEG) and Electrocardiogram (ECG). The individual results from the EEG and ECG are improved using this combination method.

EEG algorithm:

A time frequency (TF) based seizure detection algorithm is presented. The proposed technique uses features extracted from the Singular Value Decomposition (SVD) of the TF representation of EEG. These features are used with a simple Linear Discrimination Analysis (LDA) for classification of EEG traces into seizure and non seizure activity. A seizure classification accuracy was achieved outperforming most existing algorithms.

ECG algorithm:

A seizure detection technique which fully utilizes the ECG wave by extracting all the features which are found to be effected during a seizures is presented. In the previous approaches focus was only placed on the RR duration but none of the researches focused on the other

features of an ECG wave which are affected during a seizure. In our research we included RR mean, RR variance, QT duration, PR duration, P wave height and variance as the features to train Linear Discriminant Analysis (LDA). These features are found to be different for a healthy and a seizure affected individual in the literature. The results showed a classification accuracy which outperform the previous seizure detection techniques.

Combination:

Dempster Shafer rule is used for combination of the above two algorithm. The combined classification accuracy obtained outperforms any existing seizure detection algorithms.

THESIS ABSTRACT (ARABIC)

ملخص الرسالة

الاســــــــــم : محمد عبد العظيم صديقي

عنوان الرسالة : تهدف الدراسة إلى تطوير طريقة جديدة للتحليل المشترك للتخطيط الكهربائي للقلب (ECG) والتخطيطات الكهربائية للدماغ (EEG) .

التخصــــــــص: الهندسة الكهربائية

أغسطس **تاريـخ التخــرج:** 2011

تهدف الدراسة إلى تطوير طريقة جديدة للتحليل المشترك للتخطيط الكهربائي للقلب (ECG) والتخطيطات الكهربائية للدماغ (EEG) .

وتعتمد هذه الطريقة على مبادئ الأدلة النظرية لدمستر وشافر(DS).ويعتمد تحليل التخطيطات الكهربائية للدماغ (EEG) على طريقة الزمن و التردد (time-frequency) للتعرف عن النوبات القصيرة وذلك باستخراج سمات مميزة من هذا التحليل.

أما فيما يخص تحليل تخيط القلب (ECG) فنقترح استعمال طريقة المويجات (wavelets). وهذا التحليل يؤدي إلى استخراج عدة سمات نذكر منها فاصل RR، PR ، QR إلى غيرذلك.

ونذكر أن في كل من الحالتين نستعمل طريقة التحليل التميزي الخطي(LDA) لتصنيف الإشارات إلى إشارات عادية أو إشارات نوبات مرضية. وللتحسين من أداء النظام المقترح،

قدمنا طريقة مزج EEG وECG باستعمال نظرية DS والتي أدت إلى تحسين أداء النظام في تحديد الزمن والتعرف على النوبات الدماغية بنسبة تفوق 97%.

s

CHAPTER 1

INTRODUCTION

1.1 Introduction

Seizures pose a greater threat to humans with the adverse effects it can have on brain which was reported in the past. It is the most common nervous system disorder today. There are many evidences in the past related to the dangerous effects seizure can have on the normal functioning of the neurology of human beings, which may increase the risk of death[**1**][**2**]. It was found in a survey in US that almost 6% of the low birth weight infants and approximately 2% of all newborns admitted in the neonatal ICU to have seizures[**3**][**4**]. It was also found that about 2% of adults have a seizure at some time during their life[**5**]. Although there are few cases of death resulting due to seizure directly, it affects the quality of life. Upto 75% of adults with seizure were reported to have depression and are more likely to commit suicide[**6**]. The grand mal seizure if occurs during driving a cars, swimming or any such action involving continuous motion may result in an accident and ultimately to the death of an individual. Also there are many seizure which are silent in nature and if not treated may result in brain damage. Thus there is a need for detection of seizure at an early stage in order to prevent further damages to brain. The problem is that the jerky movements which are due to some other reasons may also be some time misinterpreted as seizure. This may result in the patient to receive multiple antiepileptic drugs (AEDs) over many days. The individual may become more sedated and may remain for a long time in hospital as a result of this false diagnosis. Electroencephalogram (EEG) is used as a reliable tool for detection of early seizures but the main drawback which limits the use of EEG is the lack of specialists who can correctly interpret the EEG data. Nevertheless, detection of seizure is even challenging for the neurologist by visual inspection because of myogenic artifacts[**7**]. Thus there is a need for an automatic seizure detection technique in order to reduce the false negative and false positives. Many researchers in the past have proposed Automatic seizure detection algorithms in the past based on EEG and some researchers realized the detection of seizure based on Electrocardiogram (ECG). In this work we are going to present a novel algorithm based on the combination of algorithms based on ECG and EEG.

1.2 Some Basic Definitions

Figure 1. 1: Lateral view of Brain [8]

Most common thinking when we listen to the word "seizure" is a person will shout, behave indifferently, have no control over his muscles or even lose his bladder control. This effect is just for few minutes, and the person affected with it will recover back to normal state. However this is only a form of seizure known as tonic-clonic seizure, but this is not the only kind there are several other kinds of seizure with different symptoms and in some cases no symptoms at all[**8**].

The Epileptical seizure was mentioned in the Babylonian literature 3000 years ago. The strange acts resulting from the epileptic seizure had led to various superstitious beliefs regarding epilepsy. The person undergoing seizure was thought to be possessed by demons or godly spirit. Later in 400 B.C Hippocrates, a great physician pointed out it to be a brain disorder which results when some of the neurons function abnormally.

"A seizure is the physical findings or changes in behavior that occur after an abnormal electrical activity in the brain"[**9**] . Seizures are symptoms of abnormal activity of brain resulting from abnormal firing of neurons. The function of neuron in a normal manner is responsible for the normal functioning of various glands, human thoughts & feelings. It generates electrical impulses at a rate of 80 pulses per second which moves to and fro in between the nerve cell producing different emotions, feelings and thoughts. During a seizure the neurons generate the electrical impulses at a rate of more than 500 times per second, which is very much high compared to normal rate. This causes the seizure and if the seizure occurs repeatedly it is called as epilepsy[**8**]. This can affect a part of the brain, or the whole brain depending on which it is classified into different forms of seizures. It is a sudden surge of electrical activity which leads to difference in the individual activity manifested in the form of change in perception, behavior, thinking or many times it will be hardly noticed[**10**]. It generally lasts from few seconds to maximum of about 5 minutes.

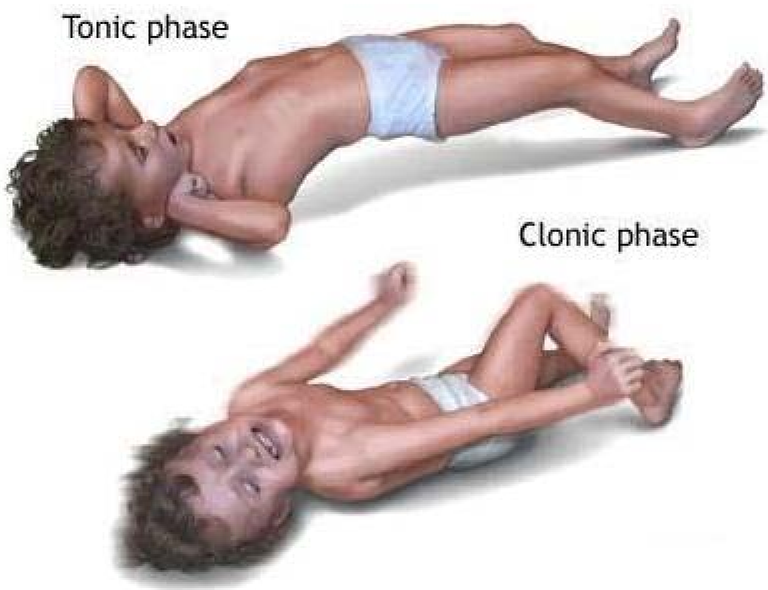

Figure 1. 2: A Boy undergoing tonic-clonic seizure [12]

The symptoms of seizures as clinical manifestation in the form of uncontrolled muscle movement, jerking are not the only real seizures but the seizure many a times result in the form of hallucination, fear, strange feeling in stomach, blanking out for a few seconds and unconsciousness which are very silent and the person does not doubt it to be a seizure[**10**]. "Symptoms of seizure occur suddenly and may last upto few minutes and may include one of the following symptoms

- Loss of control over Muscles and falling unconsciousness suddenly.
- Muscle movement such as twitching which causes the up or down motion of hand or leg.
- Tension/tightening of Muscles that causes twisting of the body, head , arms or legs

- Change in the emotional behavior. The person may experience unexplainable fear, joy or laughter.
- Changes in vision of the person. This may include hallucination or flashing of lights (seeing things that aren't there).
- Changes in sensational behavior of the skin. This may result in feeling of something spreading over the arm, body or legs.
- Changes in consciousness of the person. This may result in a person not able to have control over consciousness over some period of time.
- Change in the taste. This may be in the form of tasting something bitter or metallic flavor"[**9**]

1.3 Causes of Seizures

Seizures are linked to many reasons in the past. It happens when there is an imbalance between the neuro transmitters which help in the transmitting the electrical impulses between the nerve cells. Most researchers say it happens when there is either an abnormal increase in the neuronal activity resulting from high excitatory neurotransmitters or abnormal decrease in the neuronal activity in the brain. The most important neurotransmitter which was found to be play an active role in epilepsy was found to be gamma-aminobutyric acid (GABA) and glutamate[**11**].

"The cell membrane surrounding the neurons also plays a vital role in the seizure as the generation of electrical impulses by the neurons is dependent on them. Studies related to cell membrane such as how the molecules in the cell membrane move in and out of the membranes, and the way cell membrane nourishes or repairs the membrane reveals the fact that any hindrance in the above mentioned processes may cause the seizure. A research carried out on an animal brain showed that as the brain is adaptive to changes occurring in the stimuli continuously, if there occurs any change in the normal behavior of neuronal activity and repetition of the act may lead to a full blown epilepsy"[**11**].

About 50% of the seizures have no reason. Yet for other type of seizures they are related to one of the following problems

- Head Injury

  Head injury in some cases may lead to seizure attack although it might not be at the exact moment the injury is caused its affect may be realized at a later time[**8**].

- Heriditary Causes

  Some researchers view abnormality in a specific gene which is hereditary as one of the factor which contributes to seizure. Many seizures like progressive myoclonus epilepsy are linked to problems related to missing genes which causes a person to be susceptible to seizure activities. Dysplasia is also other kind of seizure which develops due to abnormalities in the gene structure that control neuronal migration[**8**].

- Prenatal injuries

  This occurs in the development stages of children whose brains are susceptible to many injuries like maternal infections, poor nutrition and oxygen deficiency that may harm the development of the brain of the neonates. Advanced brain imaging revealed the fact that most of the seizure cases are associated with dysplasia in the brain which are the seizures which develop before birth`[**8**].

- Environmental causes

  Mental stress, lack of proper sleep, over dosage of some drugs and exposure to carbon monoxide or other chemical may sometimes result in seizure

- Other disorders

  Seizure may develop for any event which can result in brain damage. Many diseases like brain tumors, Alzheimer's disease and alcoholism may also in some cases lead to seizures[**8**].

1.4 Different types of seizures

The Seizures are classified based on the on the part of the brain which is affected during the seizures. They are broadly classified into two types: Focal seizures and Generalized seizures.

1. Focal seizures

This occurs in about 60% of the cases of the seizures. It has an effect only on a part of the brain. It is also called as partial seizure. Depending on the area of brain which is affected it is further classified as

- Simple focal seizure

It results in unusual changes in the emotions of an individual. The individual affected with it may experience unusual joy, fear, hunger and change in emotional reactions. In some cases there are changes in the senses related to hearing, taste and seeing. The person may listen to some hallucinations, or feel the presence of someone, change in taste etc[**11**].

- Complex focal seizure

The complex focal seizure is related to the loss of consciousness , abnormal body motions, repetitive movements like walking around a circle, blinks etc. These repetitive movements are also called as automatism[**11**].

2. Generalized seizures

These seizures are results of abnormal neuronal activity resulting in all parts of the brain. This is manifested in the form of tonic-clonic seizures, tightening of arms or legs etc. The person affected may go into unconsciousness without any symptoms. The types of generalized seizures are[**11**]:

- Absence seizures
- Tonic seizures
- Clonic seizures
- Atonic seizures
- Myoclonic seizure
- Tonic-Clonic seizures (Grand mal)

The seizures can start with first being focal and then may spread to different parts of the brain resulting in generalized seizures.

1.5 Dangers of Seizures

Apart from the miscomfort caused by the seizures in day to day life of a human being there are two main life threatening conditions resulting from the seizure.

1. Status Epilepticus

Any seizure event which lasts more than 5 minutes is considered to be as Status epilepticus. A person undergoing this type of seizure will face difficulty in regaining back consciousness. “According to a survey in United States, it was found that about 60% of the people affected with it have no previous history of seizures. In United States about 42,000 deaths are noted down each year due to status epilepticus”[**8**].

2. Sudden Unexplained Death

Sudden Unexplained Death popularly known as SUDEP result due to longer Q-T duration in the ECG wave of a person during seizure. The seizure is not the only reason for SUDEP but it can increase the causes for it. This may result in a sudden death of a person without any symptom [**8**].

1.6 Problem Statement

In recent years many algorithms for detection of seizures based on electroencephalogram (EEG) have been proposed. However it was also found that in several cases, seizures are also associated with changes in heart beat rhythm and respiration rate. The affect of complex seizures can be found in the cardiovascular system and hence seizures can result as variation in the cardiac rhythm. Even though, there exists an extended body of work in the seizure detection based on ECG, much less work can be found related to the combination of the above two techniques. Previous work done related to the combination of the ECG/EEG used fusion techniques for decision making based on Bayesian formulation. However, this approach lacks in providing a meaningful solution as the Bayesian formulation of decision making assumes a Boolean phenomena which leads to over commitment i.e. the degree of belief we have in existence of certain hypothesis (say θ=Seizure). Hence a small degree of belief in a certain hypothesis θ automatically leads to large degree of belief to the negation of the hypothesis ($\bar{\theta}$). To avoid such over commitment, it is necessary to develop new approaches for fusing information from EEG and ECG without over commitment. This is exactly what we plan to investigate in this thesis. In particular, we propose to use the theory of evidence rather than the Bayes theory to fuse information from two independent classifiers, one based on EEG signal analysis and the second based on the analysis of ECG signal.

1.7 Research Objectives

The main objectives of this research are:

1) To develop an algorithm using time frequency analysis for EEG feature extraction and classification using LDA.
2) To develop an algorithm for ECG feature extraction and classification using LDA.
3) To combine the above two techniques using Dempster Shafer theory of evidence to improve classification results.

1.8 Organization of Thesis

The thesis work is organized as follows

In Chapter 2, we will be discussing the literature review related to the various seizure detection techniques proposed in the past based on Electroencephalogram (EEG), Electroencephalogram (ECG) and other techniques. A literature review of different combination methods for the seizure detection techniques used in the past will also be discussed in this chapter.

In Chapter 3, we propose a seizure detection technique which is based on time frequency approach of EEG signal. The left singular vector of the time frequency matrix of EEG signal is used as feature vector to train linear discriminant network to classify the results as seizure and non seizure.

In Chapter 4, we propose another seizure detection technique which is based on features extracted from ECG signal. The features extracted are again fed to linear discrimination analysis for classification.

In Chapter 5, we propose to combine the results obtained in Chapter 3 and Chapter 4 using Dempster Shafer theory of evidence (DST). The reason for using DST and conceptual difference between the Bayesian theory and DST are discussed.

In Chapter 6, we conclude the thesis by making some concluding remarks and mentioning the scope for future work on this topic.

1.9 Section Summary

In this section we have discussed the concept of seizure and different types of seizures. We have also discussed the effect of these seizures on human being and the threat posed by seizures to an individual’s life. The need for seizure detection techniques at an early stage may help in reducing the risk of life posed by seizures. For achieving this we have proposed a new seizure detection algorithm which can detect seizures more accurately, so that the issue can be handled before time. Finally, we have discussed the main objectives of our thesis and strategy for achieving the goals in the further chapters.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This section discusses the literature survey of various papers done in order to understand the research work done by other researchers in similar field. The detection of seizures is generally based on the processing of signal data from brain. But in the past seizure detection algorithms were presented which were dependent on the processing of the signals from heart and other body movement. In the following sections, we are going to discuss the various algorithms dependent on various signals from the body used for detection of seizures in the past.

2.2 Biomedical Signal Processing

In recent years biomedical signal processing has gained very much popularity for its contribution in the field of medical sciences. It is used in extracting information related to various physiological activities varying from protein and gene sequences, to neural and cardiac rythms to tissue and organ images[**12**].

In the past, research was focused on filtering biomedical signals to remove the artifacts and noise. The noise is generated in capturing signals from different parts of the body due to the instrument contacts, precision, and the biological system under study. Removing the unwanted noise can reveal the information underlying. Different approaches are used for removing the noise. Apart from these noise cancellation techniques, many biomedical instruments are developed for analyzing biological signals.

“The use of biomedical signal processing in the present is focused on the medical imaging modalities such as ultrasound, Magnetic Resonance & Imaging (MRI), and positron emission tomography (PET). It enables radiologists to visualize the structure and function of human organs. Cellular imaging such as fluorescence tagging and cellular MRI assists biologists in monitoring the distribution and evolution of live cells; tracking of cellular motion and supports modeling cytodynamics. The automation of DNA sequencing aids geneticists to map DNA sequences in chromosomes. Analysis of DNA sequences extracts genomic information of organisms. The invention of gene chips enables physicians to measure the expressions of thousands of genes from few blood drops. A Correlation study between expression levels and phenotypes unravels the functions of genes”[**12**]. The above examples show that the signal processing made a great contribution in the field of biomedicine.

2.3 Seizure detection based on Electroencephalogram (EEG)

“Electroencephalography (EEG) is the recording of electrical activity along the scalp produced by the firing of neurons within the brain”[**13**]. In clinical terminology, it means the recording of activity of brain over a time period. This is an important tool in detecting early seizures. Many studies have reported dealing with the automatic detection of seizures based on EEG in the past.

A.Liu et al [**14**] shows that the periodicity and autocorrelation analysis of the EEG signal as the dominant characteristics of seizure and used autocorrelation analysis to quantify rythmicity in EEG. It was observed that the electrographic seizures are generally silent in nature and were distinct from the normal background cerebral activity. The autocorrelation analysis is hence used to distinguish the background cerebral activity from the seizures. The autocorrelation of a seizure pattern was shown to consist of peaks regularly spaced with same frequency as the original signal whereas for a non seizure trace it showed to consists of irregular spaced peaks and troughs and hence it is easy to detect the seizure pattern from the non seizure based on this spacing. This method popularly known as Scored Autocorrelation Anlayis (SAM) was found to give a sensitivity of 84% and specificity of 98%. This is the first attempt of seizure detection using EEG and the results obtained are quite good. This is the first method which provided an idea for the researchers to dwell into the area of automatic seizure detection using EEG.

J.Gotman et al[**15**] used a combination of automated methods too increase the detection rates and decrease the false alarms. They discussed three different methods for the analysis of the EEG signal. The 3 different methods are: 1) Spectral analysis for detection of rhythmic discharges at various frequencies; 2) Spike detection for finding group of signals which do not have rhythmic nature and give abnormal spikes instead; 3) Low pass digitally filtered EEG signal for finding very slow discharges. For the spectral analysis the authors have used the Fast Fourier Transform (FFT) based frequency spectrum analysis to detect periodic discharges. The frequency spectrum of each 10 sec epoch is calculated and a number of features such as frequency , width of the dominant spectral peak, and relative power of frequency bands were extracted. The spike detection of the EEG trace is performed by passing the given EEG trace through a high pass filter. The detection of very slow rhythmic discharges is performed by passing the signal through a low pass filter. The algorithm was able to detect 71% of seizures and 78% of seizure clusters were detected with a false detection rate of 1.7/h.

In another evaluation technique carried out by J.Gotmal et al[**15**] on various data provided by three different institution from Canada, the USA and Australia showed a detection rate of 77%, 53% and 84% respectively.

Osorio I et al [**16**] developed an algorithm which uses time frequency localization, signal processing, and identification of time frequency stochastic systems to detect seizures. The algorithm was able to detect 92% of the seizures accurately.

P.Celka and Paul Colditz [**17**] proposed a SSA-MDL (Singular Specturm Analysis- Minimum Description Length) based algorithm for detection of seizures. The author based the algorithm on the fact that the seizure has an effect of producing synchronous discharge (rhythmical activity) of neurons whereas a non seizure activity has asynchronous discharge of neurons (non rhythmical activity). As the Singular Spectrum Analysis is found to have given good results in biomedical signal processing application Singular Value Decompostion is used for analysis of EEG signal. The second part of the algorithm is to find the optimal dimension estimation n_o which is found using the Rissanen's Minimum Description Length criterion. The n_o is very important as it decides the amount of stochastic content in the EEG signal. The value of $n_o \approx 3$ is used to prove that the signal was originated from a low dimension system, which can be used for detection of rhythmic activity. The algorithm showed a good detection rate of 93% and false detection rate of less than 4%. The algorithm requires a lot of computational load and increases the time of computational execution.

P.E.McSharry et al [**18**] proposed a non linear technique which uses Multi dimensional probability evolution (MDPE) which can detect the underlying dynamics related to EEG. The authors compared the variance based seizure detection technique with the non linear analysis of the EEG signal for 10 EEG traces and found that the non linear analysis gives fewer false positives compared to variance based analysis but no firm belief is established that the MDPE can outperform the variance based method in identifying seizures.

Reza Tafreshi et al [**19**] proposed a wavelet based method for detection of seizures with temporal lobe epilepsy. The detection method identify the nodes of a wavelet packet by using the local discriminant bases and cross data entropy algorithms. Based on the results obtained with the limited data they have, the authors concluded that wavelet packet energy ratio could be used as a good criterion for classification of seizure and non seizure patterns.

N.Kannathal et al [**20**] proposed the use of different entropy estimators for distinguishing a healthy EEG trace from a seizure one. It was found to give an accuracy of 90%.

Abdulhamit Subasi [**21**] proposed a neural network based approach which uses Dynamic fuzzy neural network (DFNN) for classification purpose. The EEG signal was first decomposed using discrete wavelet transform of level 5 into different frequency sub bands. These wavelet coefficients were used for training the DFNN network. The results showed an accuracy of 93% with a specificity and sensitivity of 92.8% and 93.1%.

H.Hassanpour et al [**22**][**23**] proposed a time frequency based feature extraction algorithm. The technique used the left and right singular vectors of the time frequency distribution of the EEG signal to differentiate between a seizure and non seizure activity. The estimated distribution function related to seizure and non seizure epochs are used to train a neural network to discriminated between seizure and non seizure patterns. The results showed 90% and 5.7% good detection rate and false detection rate respectively. The false detection rate is more in this case which can result in false detection of seizures in healthy cases. A more improved version of this can be deemed to be usable in real time seizure detection.

Hojjat Adeli et al [**24**] presented a Wavelet-Chaos methodology. The technique uses correlation dimension (CD) and largest Lyapunov exponent (LLE) which represents system complexity and chaoticity are used for differentiating healthy and epileptic traces. The EEG signal is decomposed into different frequency bands named alpha, beta, theta, gamma and delta by wavelet decomposition. The Correlation dimension (CD) and largest lyaponov exponent (LLE) are calculated for each sub band and are used for differentiating between the seizures and non seizure event. It was found that for higher frequency sub bands like beta and gamma, Correlation dimension (CD) effectively differentiates between the seizure and non seizure trace, whereas for lower frequency bands like alpha LLE effectively differentiates between the seizure and non seizure traces. The author discussed presented in this case a new method for seizure detection but nothing was done experimentally on the EEG data.

Ardalan Aarabi et al [**25**] developed a seizure detection technique where the features extracted from the EEG signal are selected through relevance and redundancy analysis. The extracted features are then trained using multilayer back-propagation neural network. The classification resulted in an accuracy of 79.7% detection rate with a sensitivity and selectivity of 74.1% and 70.1%.

Bedakh Abibullaev et al [**26**] propsed a seizure detection method based on the best basis wavelet functions and double thresholding. The algorithm first decomposes the

EEG trace with the wellknown wavelet functions such as Daubechies family db2, db5 and from the biorthogonal family bior 1.3, bior 1.5 and then applying thresholding for denoising and classifying the EEG traces into seizure ictal and interictal states. The results showed a Good detection rate and False detection rate of 93.2% and 5.25% respectively for seizure events and 90.75% and 8.25% for seizure interictal events.

Anup Kumar Kesri et al [**27**] presented a Epileptic spike detection technique which uses Deterministic Fintie Automata (DFA) for finding the spikes in a EEG seizure trace. With 10 EEG signal data the recognition rate was found to be 95.68%.

Zandi AS et al [**28**] proposed a wavelet based algorithm which uses wavelet coefficients from seizure and non seizure to differentiate between seizure and non seizure. A Combined seizure index (CSI) is developed by representing the separation between the seizure and non seizure states in frequency bands. CSI is derived for each EEG trace of seizure and non seizure states based on the rythmicity and relative energy. The results showed a sensitivity of 90.5% with false detection rate of 0.51 h-1.

Apart from these many techniques were presented in the past [**29**] [**30**][**31**][**32**]. Those mentioned here are the major works related to detection of seizures using EEG.

2.4 Seizure detection based on Electrocardiogram (ECG)

Less research is done in the field of seizure detection using ECG signal. Here, we are going to present the work of previous researchers on detection of seizure using ECG signal.

D.H.Kerem and A.B.Geva [**33**] have proposed an algorithm which proposes to use the information contained in RR-interval series which includes the R-R interval duration and differential R-R interval with respect to the previous R-R duration and applied to an unsupervised fuzzy clustering algorithm which rendered them with a success rate of 86%. This method uses only the RR information for seizure detection and nothing has been mentioned related to other features of ECG signal.

Barry R.Greene et al [**34**] proposed a linear discriminant classifier which processes 41 heartbeat timing interval features. The features used in this study included: mean RR interval, relative mean RR interval, RR interval standard deviation, the relative mean standard deviation, RR interval coefficient of variation, RR interval power spectral density (PSD), change in RR interval, relative change in RR interval, RR interval spectral entropy. The method came up with an average accuracy of 70.5% and associated sensitivity of 62.2% and specificity of 71.8% for a patient specific basis. On a patient independent basis it achieved an accuracy of 68.3% with a sensitivity of 54.6% and speicificity of 77.3%. Here also the algorithm came with different features related to RR interval and the accuracies obtained are very less compared to other available techniques.

M.B.Malarvili et al [**35**] proposed a Heart Rate Variability (HRV) as a tool for assessing seizure detection instead of seizure detection instead of R-R interval. The time frequency distribution of HRV is obtained and features related to mean and variance of HRV in low frequency band (0.03-0.07 Hz), mid frequency band (0.07-0.15 Hz), and high frequency band (0.15-0.6 Hz) are used to discriminated between a neonatal seizure from the non seizure. The technique was found to give a maximum of 83.3% of sensitivity and 100% specificity. The authors presented the algorithm without performing any test on real time ECG data.

M.B Malarvili and Mostefa [**36**] proposed to use both the features in time domain and time frequency domain of R-R interval and Heart Rate Variability (HRV). The time domain features include mean and standard deviation of RR interval and Hjorth parameters, which describe the characteristic of a signal in terms of activity, mobility, and complexity were computed for HRV. The time frequency distribution includes mean, standard deviation, rms, min, max , coefficient of variation, skewness, and kurtosis of the intermediate frequency (IF), Intermediate Bandwidth (IB) and energy in LF, MF, and HF, the total energy in all HRV components and the ratio of energy concentrated in the LF to HF (LF/HF) were considered. Finally, the features from both time domain and frequency domain were selected and optimal features were used for classification of signals.

In all the above techniques it was observed that the only focus made in the seizure detection algorithms related to ECG signal is on the RR interval and no research is done on the other features related to ECG signal such as PQRST waves of ECG and their sub features.

2.5 Seizure Detection Based on Other Methods

Apart from the use of ECG or EEG seizure detection based on body movement was also proposed. A seizure detection algorithm based on Electrocorticography (ECoG) was also presented by the researchers. In Electrocorticography (ECoG) the electrical activity of brain is recorded directly by placing the electrodes over the surface of brain from the cerebral cortex. It is known to be "gold standard" for detecting seizure in clinical practice. This is done during the surgery or outside the surgery in Intensive Care Units[**37**]. Based on the usage of ECoG Osorio I et al [**38**] proposed a real time seizure detection algorithm which is based on wavelet decomposition of the ECoG trace. The testing was performed with 14 subjects and results showed a sensitivity of 100% without adaptation. After adaptation 2 undetected seizures and two unclassified seizures were captured.

N.Karyiannis et al [**39**] proposed a new seizure detection technique which depends on the body movements of the neonates rather than EEG/ECG recordings. This method depends on the body part movements of the neonates recorded through standard video recorders. The authors used image segmentation and motion tracking to quantify neonatal movements in the video recordings of 54 neonates with seizures. The results provided an effective strategy for training a neural network to automatically recognize neonatal seizures. The major drawback of this method is that it does not utilize EEG and therefore cannot detect vast majority of neonatal seizures i.e purely electrographic or subtle seizures.

2.6 Combination of Seizure Detection Algorithm

In medical decision making biomedical data fusion consists of combining data, reducing its complexity and designing a synthetic representation to be more easily interpreted. This requires the integration of seizure detection techniques to give good results. The different types of fusion techniques can be thus classified as follows:

A. Classification based on feature combination

   The first type of classification is based on the method of combination of features from the different seizure detection algorithm. They are classified into two types:

1. Early fusion of features:

This type of fusion technique involves concatenating the EEG and ECG feature vectors into a single feature vector and feeding this ‘super vector’ to a pattern classifier as illustrated in figure 2.1.

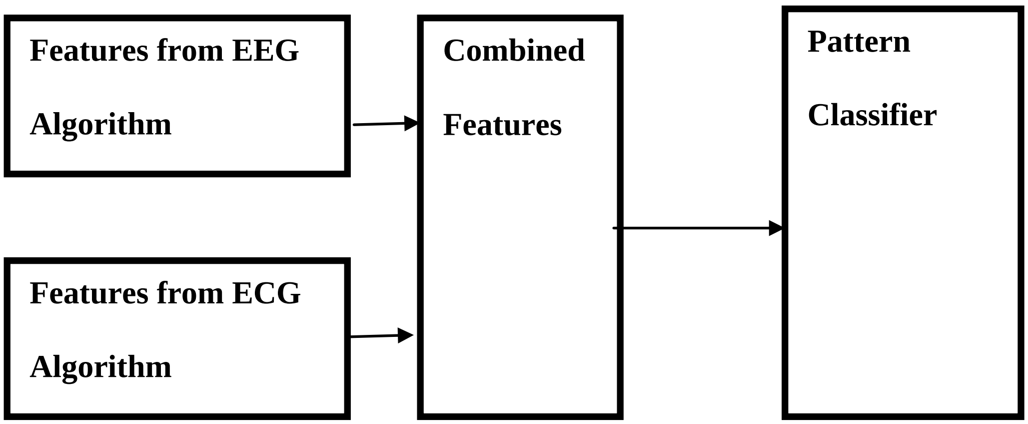

Figure 2. 1: Early fusion of features

2. Late fusion of features:

This type of fusion technique employs separate classifiers for each signal to determine a probability of seizure for each signal mode. These two probabilities are then combined to give an overall probability of seizure as shown in figure 2.2. Based on the combined probability the decision is made.

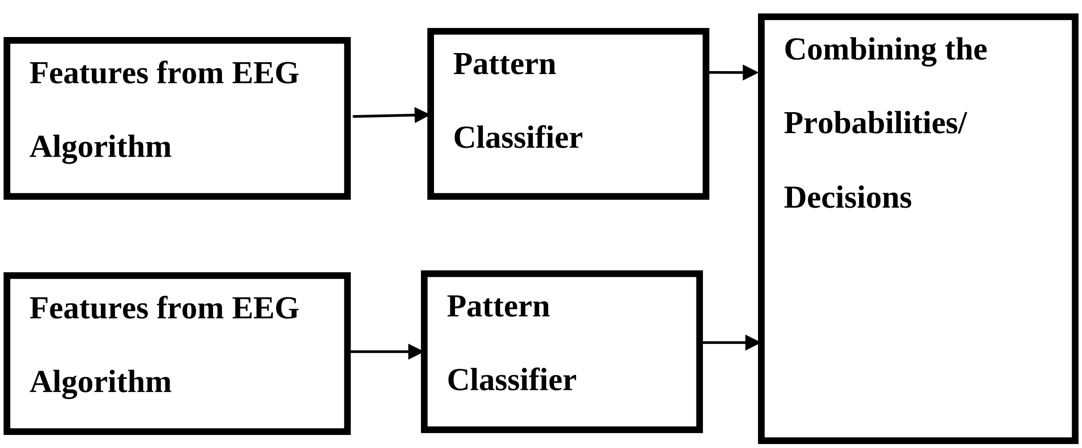

Figure 2. 2: Late fusion of features

B. Classification based on decision making

The second type of classification is based on the method of decision making which is classified into two types:

1. Fusion of probabilities

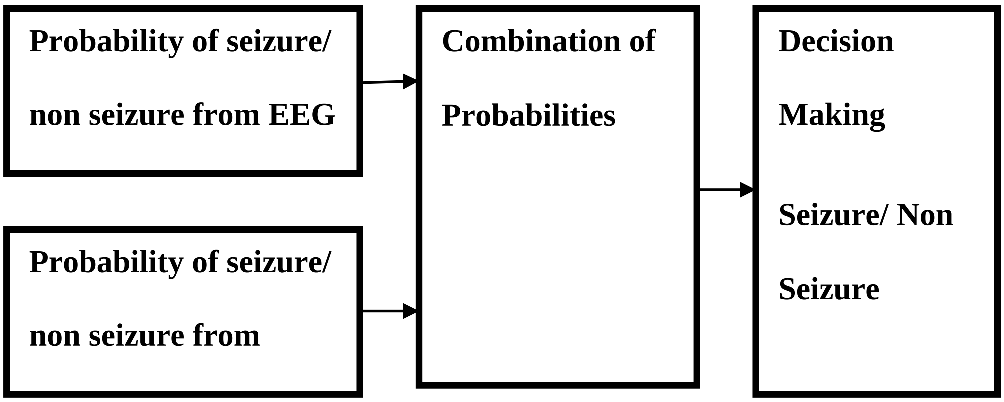

Figure 2.3: Fusion of probabilities

In this intermediate scheme the feature vectors are reduced to probability vectors which are fused in a common global fusion centre as illustrated in figure 2.3.

2. Fusion of decisions

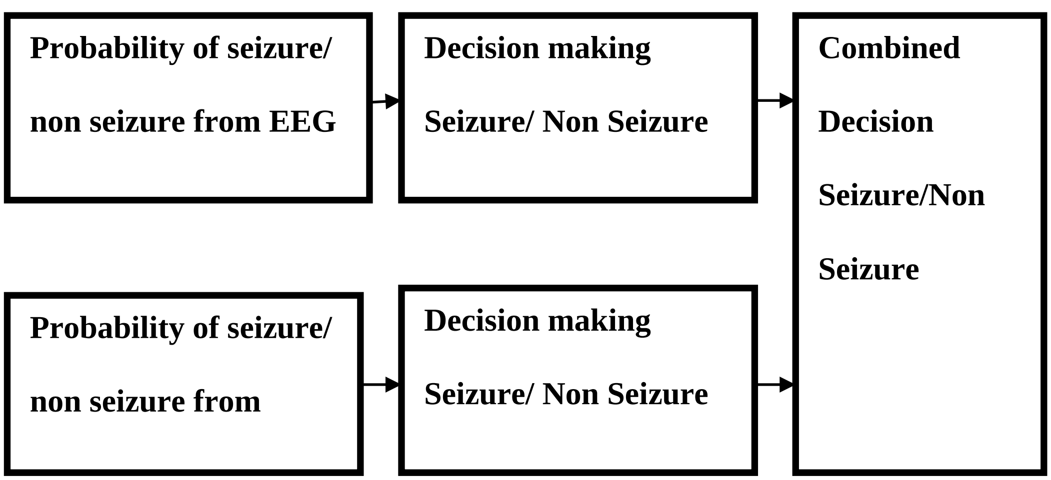

Figure 2.4: Fusion of decisions

In the technique illustrated in figure 2.4 the feature vectors are reduced to probability vectors through their own forecaster. The partial decisions made by the decision makers based on the probabilities are fused through a global decision maker. In this scheme, the partial decisions are set to 1 when the posterior probability of the corresponding modality of data is greater than 0.5. The global decision support seizure when both partial decisions agree.

To improve the accuracy of seizure detection algorithm and to reduce the false alarms, a combination of features extracted from only EEG or ECG were introduced. Barry R.Greene et al [**40**] first attempted to improve seizure detection was made by combining EEG and ECG data simultaneously. The authors proposed two methods for fusion of data. The first method was to combine the features of both ECG and EEG together and then train the neural network with the combined features. The second method was to employ separate classifiers for ECG and EEG to determine probability of seizure for each signal mode. These two probabilities are then combined to give an overall probability of events. The first method provided a better performance compared to the later one.

T.Bermudez et al [**41**][**42**] introduced different methods for combination of EEG and ECG features. The different fusion techniques presented are fusion of features, fusion of probabilities and fusion of decisions. In fusion of features, the features of both EEG and ECG are concatenated and then fed to a classifier which gives the probability of seizure. This probability is used for decision making. In fusion of probabilities, the feature vectors are reduced to probability vectors and these probability vectors are combined. This gives an overall probability of seizure which is used for decision making.

In fusion of decisions, the ECG and EEG automatic seizure detection technique are used separately and the partial decisions made by the individual decision makers, which are based on the probabilities are fused together through a global decision maker. The global decision maker makes the decision in favor of seizure when both partial decisions agree.

2.7 Section Summary

In this section, a literature review of the previous techniques for seizure detection was presented. We discussed algorithms for seizure detection using EEG , ECG, ECoG and video recording of body movement. It was found that much research is based on the detection of seizure using EEG and fewer algorithms are proposed based on other methods. Various combination techniques possible for combining the results from various classifiers are also discussed and a literature review of combined classifiers for seizure detection is also presented. In the following chapter we will be discussing the detection of seizure based on Electroencephalogram (EEG).

CHAPTER 3

SEIZURE DETECTION BASED ON EEG SIGNAL

3.1 Introduction

An EEG trace can be seen as a summary recording of electrical activity of several billions of neurons over time along the scalp. The electric potential produced by single neurons are far too small to be recorded and hence the EEG activity therefore represents the summation of synchronous activity of neurons in similar orientation[**43**][**44**]. A standard EEG recording technique using 10-20 electrode system is shown in figure 3.1.

EEG traces play an important role in the detection of disorders related to brain. EEG is used as the main diagnostic tool for detecting abnormalities related to epileptic activity[**45**]. Its secondary applications find clinical use in diagnosis of encephalopathies, coma and brain death. It is also used to identify other problems related to sleeping disorder and changes in behavior etc.

In this thesis, we propose to use a hybrid time-frequency based linear discriminant analysis (TF-LDA) of EEG for seizure detection. It was showed, in previous research that the seizures have signatures in both low and high frequencies. It was also shown that seizure activity is best recorded in the delta range (up to 4 Hz) of EEG and also it has some signatures in the theta (4-7 Hz) and alpha ranges (8-12 Hz)[**2**]. We decided here to focus our research on the analysis of these low frequency content of EEG traces.

3.2 EEG Data

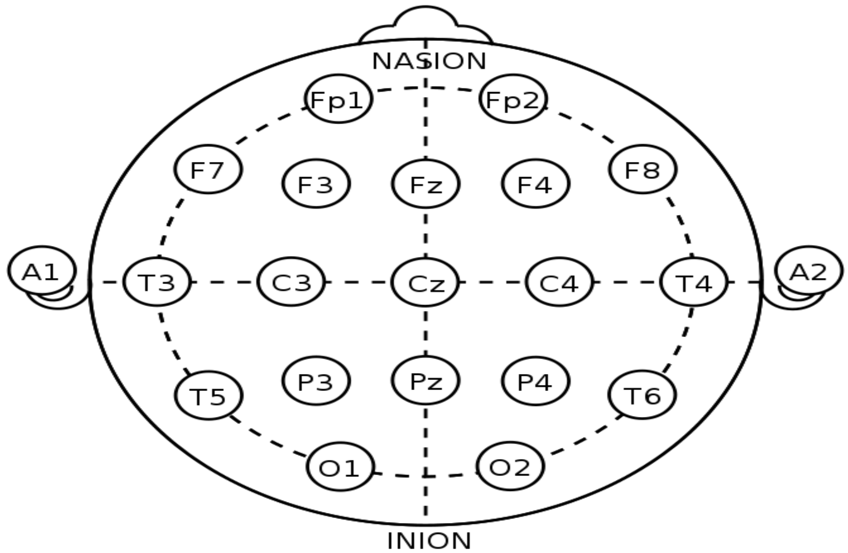

Figure 3. 1: Standard 10-20 electrode for recording [**46**]

The EEG data used in this research is provided by Dr. Ralph Andrzejak of the Epilepsy center at the University of Germany and is made available online by the authors at *http://www.meb.unibonn.de/epileptologie/science/physik/eegdata.html*[**47**]. The EEG data is recorded using the standard 10-20 electrode system as shown in the figure 3.1 [**46**]. EEG data from three different categories is presented: 1) Healthy, 2) Epileptic subjects during seizure-free intervals, and 3) Epileptic subjects during seizures. Five sets (denoted S, Z, E, F, O) each containing 100 single channel EEG segments of 23.6-sec duration, were used for our study.

The data was recorded with a band pass pre filtering of 0.53-40 Hz. The different segments were selected and cut out from continuous multichannel EEG recordings after visual inspection for artifacts, e.g., due to muscle activity or eye movements. Volunteers were relaxed in an awake state with eyes open (Z) and eyes closed (O), respectively.

Segments in sets E and F correspond to seizure free intervals, and set S is the only set corresponding to epilepsy-prone subjects during seizure. The data made available by the authors is free from any artefacts and can be readily used for further processing [**47**].

For our study, we use set Z to represent healthy subjects data and set S as the epileptic subject data. The type of epilepsy was diagnosed as temporal lobe epilepsy with the epileptogenic focus being the hippocampal formation. Each data segment contains N=4097 data points collected at 174 Hz sampling rate . Each EEG segment is considered as a separate EEG signal resulting in 200 EEG signals, 100 for healthy subjects and 100 for epileptic subjects during seizure. Two typical sample segments are displayed in figure 3.2. In the section below we are going to discuss the nature of EEG trace and the algorithm to extract the feature vector from the EEG trace.

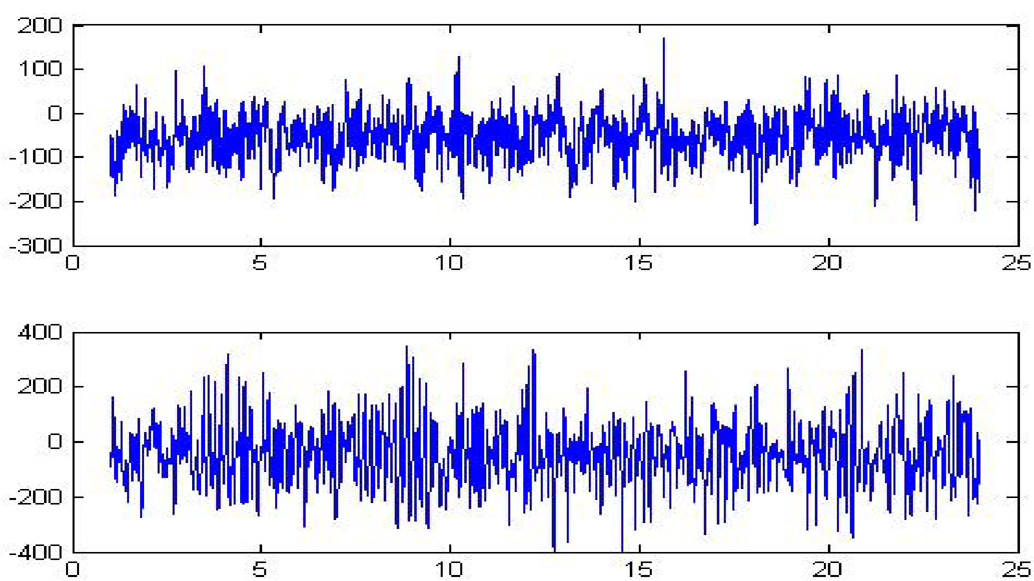

Figure 3. 2: Sample EEG signals for non seizure (top) and seizure traces (bottom)

3.3 Type and Nature of EEG trace

The type and nature of biomedical data often indicates health status of the patient. It is necessary to know the nature of signal in order to preprocess the signal for further analysis and tests to be performed.

The EEG traces, either it is recorded for a healthy person or an epileptic seizure patient were found to be non linear in their nature. The authors Ye Yuan Yue Li et al[**48**] performed a detailed research on different types of EEG traces from the dataset used in our research and concluded that the EEG traces are non linear and stochastic. It was also found that the amount of non linearity found in the seizure EEG trace is more compared to healthy EEG trace[**48**]. Earlier work on EEG signals has also shown that such signals exhibit stochastic and non stationary behavior, which means the frequency information of the signal varies with time [**49**]. Hence, the information content in the signal can't be captured either by time analysis techniques or by frequency domains approaches (such as the Fourier transform). For this reason Time frequency Represenation (TFR) techniques are used to represent the variation of frequency content of the signal with respect to time.

In clinical practice, EEG traces are usually displayed on special paper or more commonly on PC monitors. Unfortunately, time domain representation of EEG signals fail to reveal some important changes in the EEG traces easily leading to misinterpretation of EEG traces and even more seriously missing possible signs of epilepsy. For this reason, we decided to use different time frequency representation (TFR) to analyze EEG traces. In the following section, we are going to analyze which time frequency representation suits best for the representation of seizure traces.

3.4 Time Frequency Representation (TFR)

The EEG signal available in raw form, as shown in the figure 3.2 does not show any information related to the frequency content of the signal. In order to get information from non stationary signals like EEG, we need to use time frequency representation. It is well known that the time frequency representations cannot necessarily give high resolution in both time and frequency domains at the same time. The selection of a particular time frequency representation depends on the kind of application and features of interest. For this purpose, we are going to discuss below the different TF models used in the literature and test their appropriateness in modeling the EEG.

3.4.1 Short Time Fourier Transform (STFT)

The STFT is a windowed version of the Fourier transform, where the Fourier transform of a signal is taken while sliding the window along the time axis. The main disadvantage of using a Fourier transform is that it does not give any information related to the time at which the frequency component occurs. This creates a problem for analyzing a non stationary signal which consists of multiple frequency components occurring at different time. This drawback in Fourier transform is overcome by using STFT, where a moving window of fixed length is applied to the signal and Fourier transform is applied to the moving window. It is used for linear signals and is used to determine the sinusoidal frequency and phase content of local sections of signals as it changes along the time axis. The STFT of a signal x(t) is given by

$$X(t,f) = \frac{1}{\sqrt{2\pi}} \iint_{-\infty}^{\infty} x(\tau)h(\tau - t)e^{-j2\pi v\tau}\, d\tau.df \qquad (3.\ 1)$$

Where,

$X(t, f)$ is the STFT of x(t) which is the Fourier transform of the input signal x(t)

τ is the time difference between the actual signal and the shifted version

f is the Frequency

$h(\tau)$ is the windowing function

The STFT of a seizure EEG trace with different window sizes are shown in the figures 3.3 - 3.5. It can be seen from the figure that the STFT with a window size of 500 bins gives better resolution in both time and frequency compared to others.

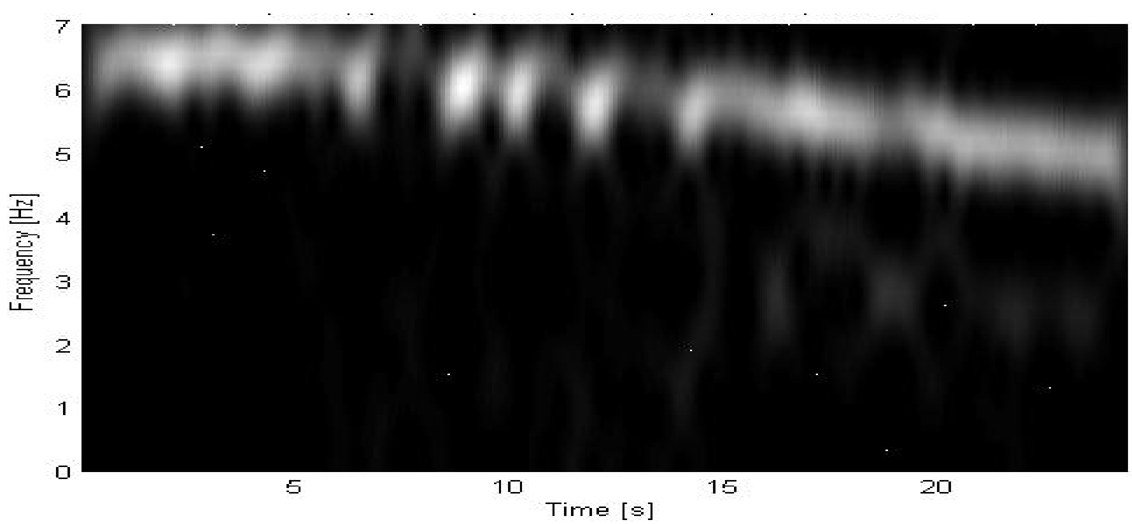

Figure 3. 3: STFT of seizure trace with a window size of 150 bins

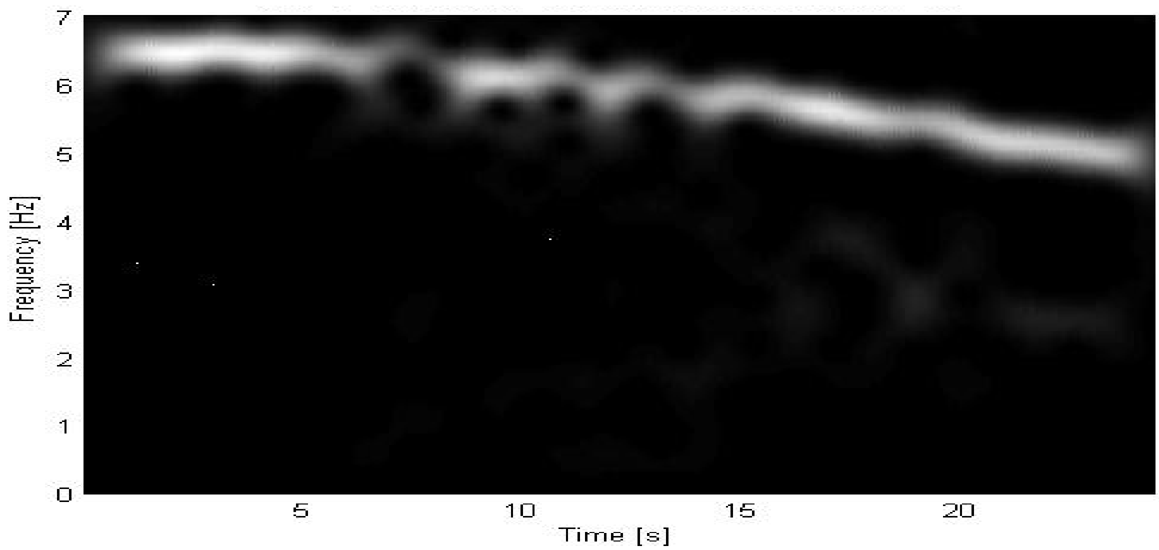

Figure 3. 4:STFT of EEG seizure trace with a window of size 300 bins

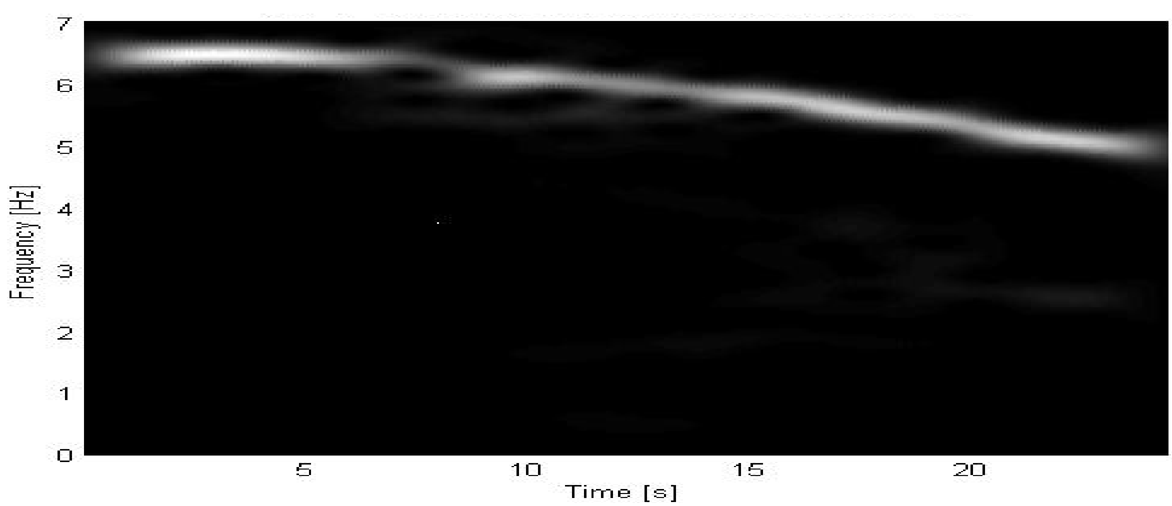

Figure 3. 5: STFT of EEG seizure trace with a window of size 500 bins

The drawback of STFT is the use of fixed window size which results in a tradeoff between time and frequency resolution. A large window will provide good resolution in frequency domain but poor resolution in time domain and vice versa. The STFT is generally used in audio signal processing applications for equalization or tuning audio effects etc.

3.4.2 Wigner Ville Distribution (WVD)

Wigner Ville distribution was introduced in the year 1932 by Wigner & Ville. It gained popularity as it is very simple found and overcame the problem of fixed window size found in STFT. It gives a better time and frequency resolution compared to STFT and hence widely used in signal analysis and has a wide range of application in signal processing, speech processing, EEGs, ECGs ,to listen heart and muscle joint sounds etc[**50**].

To overcome the problems found in the previous time frequency distribution, another method of analyzing non stationary signals was proposed. This was to perform signal analysis of Fourier transform of auto correlation function. According to Wiener – Khinchin the signal's energy of a signal $x(t)$ in time frequency domain can be considered as the Fourier transform of auto correlation function given by

$$P(t,f) = \int R(\tau)\exp(-j2\pi f\tau)\, d\tau \qquad (3.2)$$

Where,

f represents the Frequency

τ represents the time lag

And $R(\tau)$ is the autocorrelation function given by

$$R(\tau) = \int x(t).x^{*}(t-\tau)\, d\tau \qquad (3.3)$$

Where $x^{*}(t-\tau)$ is the rotated and time shifted version of the original signal $x(t)$

To make the above equation time dependent the auto correlation function is made time dependent. The time function of the equation is thus written as

$$P(t,f) = \int R(t,\tau)\exp(-j2\pi f\tau)\, d\tau \tag{3.4}$$

For Wigner Ville distribution the auto correlation is chosen to be

$$R(t,\tau) = x\left(t+\frac{\tau}{2}\right).\,x^*\left(t+\frac{\tau}{2}\right) \tag{3.5}$$

By Substituting the equation 3.5 in equation 3.2 we get

$$WVD(t,f) = \int x\left(t+\frac{\tau}{2}\right).\,x^*\left(t+\frac{\tau}{2}\right).\exp(-j2\pi f\tau)d\tau \tag{3.6}$$

The Wigner ville distributions for a seizure EEG trace with different window sizes are shown in the figures 3.6 – 3.8.

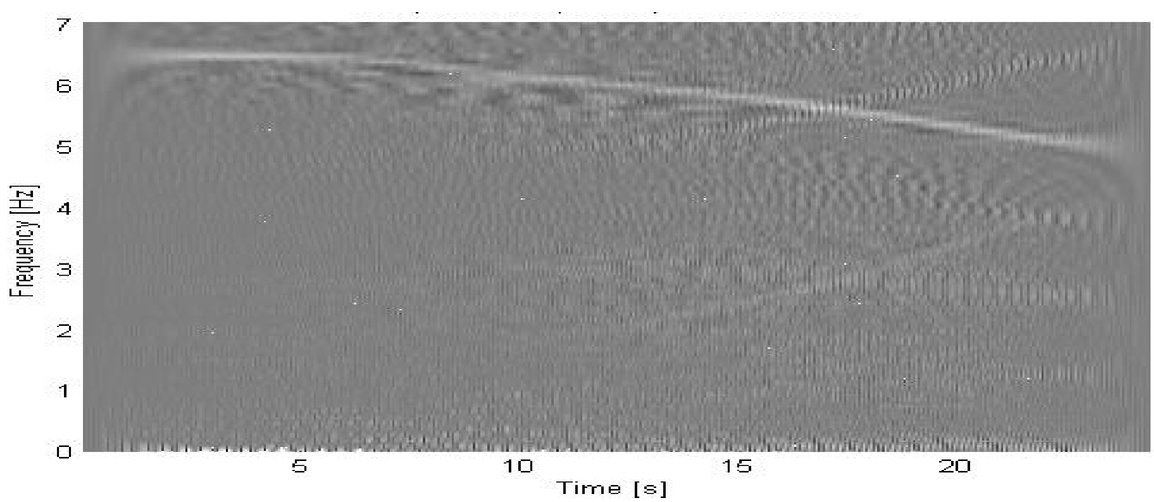

Figure 3. 6:Wigner Ville TFR for EEG seizure trace with a window of size 150 bins

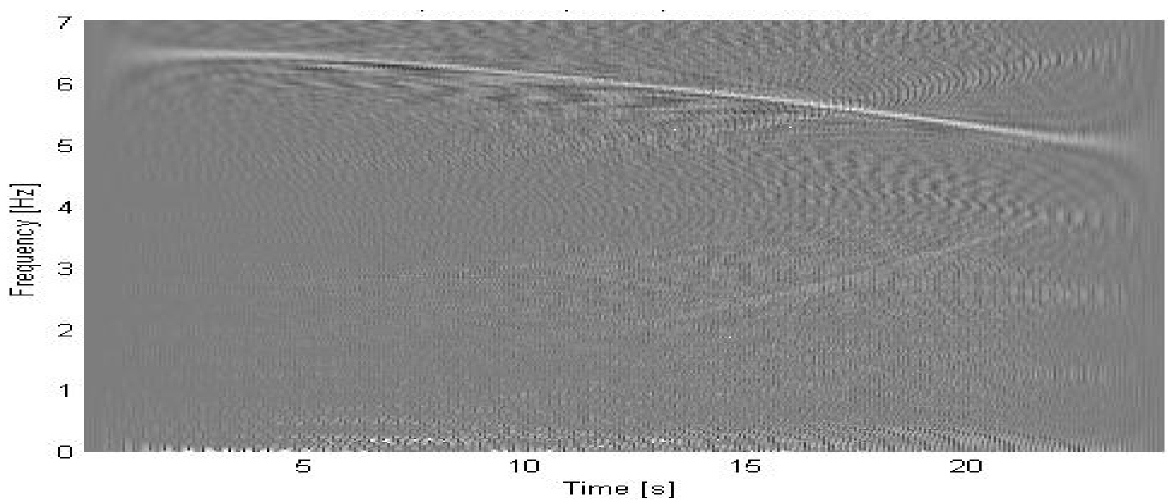

Figure 3.7:Wigner Ville TFR for EEG seizure trace with a window of size 300 bins

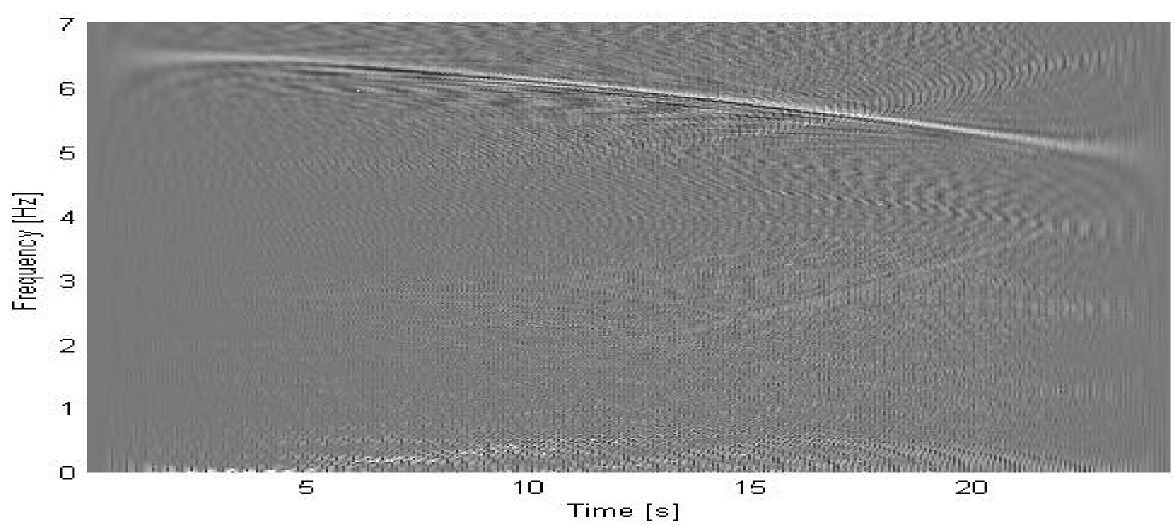

Figure 3.8:Wigner Ville TFR for EEG seizure trace with a window of size 500 bins

It can be seen from the figures 3.6 – 3.8 that the Wigner Ville distribution with a window size of 500 gives a better representation of seizure event compared to other Wigner Ville distribution. The major drawback of Wigner Ville is the introduction of cross terms which increases the interference. To reduce these cross terms other TF methods were introduced. In the next section, we are going to discuss two of the major

TF methods used for reduction of cross terms in order to have a better view of seizure events in the EEG trace.

3.4.3 Choi Williams Distribution

Choi Williams and ZAM belongs to Cohen's class of time frequency distribution. According to Cohen all bilinear TF representation can be represented in a general form [**51**]. If the Fourier transform in the equation is done with respect to t instead of τ then we obtain a popular joint time frequency distribution called as ambiguity function (AF) given by

$$AF(\vartheta,\tau) = \int x\left(t+\frac{\tau}{2}\right).x^*\left(t+\frac{\tau}{2}\right).\exp(-j\vartheta\tau)dt \qquad (3.7)$$

Where

τ is time shift

ϑ is frequency shift

Based on this AF Cohen proposed a time dependent auto correlation function defined by

$$R(t,\tau) = \frac{1}{2\pi}\int AF(\vartheta,\tau).\varphi(\vartheta,\tau).\exp(j\vartheta\tau)d\vartheta \qquad (3.8)$$

Where AF is the Ambiguity function defined in equation 3.7

And $\varphi(\vartheta,\tau)$ is called the kernel function

Cohen reduced the work for design of time frequency distribution by introducing the kernel function. Instead of designing a new time frequency distribution the researchers focused on the selection of kernel function. Based on different kernel function there are dozens of time frequency distribution proposed. One of them with a major significance is Choi Williams distribution.

Choi Williams distribution was proposed by H.Choi and W.J.Williams in 1989 to improve the time frequency representation by reducing the cross term interference [**52**]. The authors proposed an exponential kernel to the Cohens class for suppressing the cross terms. The representation of Choi Williams distribution is defined as

$$CW(t,f) = \iint_{-\infty}^{\infty} A(\vartheta,\tau).\,\varphi(\vartheta,\tau).\,\exp\big(j2\pi(\vartheta t - \tau f)\big)\,d\vartheta d\tau \qquad (3.9)$$

Where $A(\vartheta,\tau)$ is the ambiguity function given in equation 3.7 and the kernel $\varphi(\vartheta,\tau)$ for Choi Williams is given by

$$\varphi(\vartheta,\tau) = \exp[-\alpha\,\vartheta\tau^2] \qquad (3.10)$$

The larger the parameter α, the more the cross terms are suppressed. On the contrary the auto terms are increased with an increase in α. So there is a trade off between the cross terms and auto terms. The Choi Williams representations for EEG seizure trace with different window sizes are shown in the figures 3.9 – 3.11.

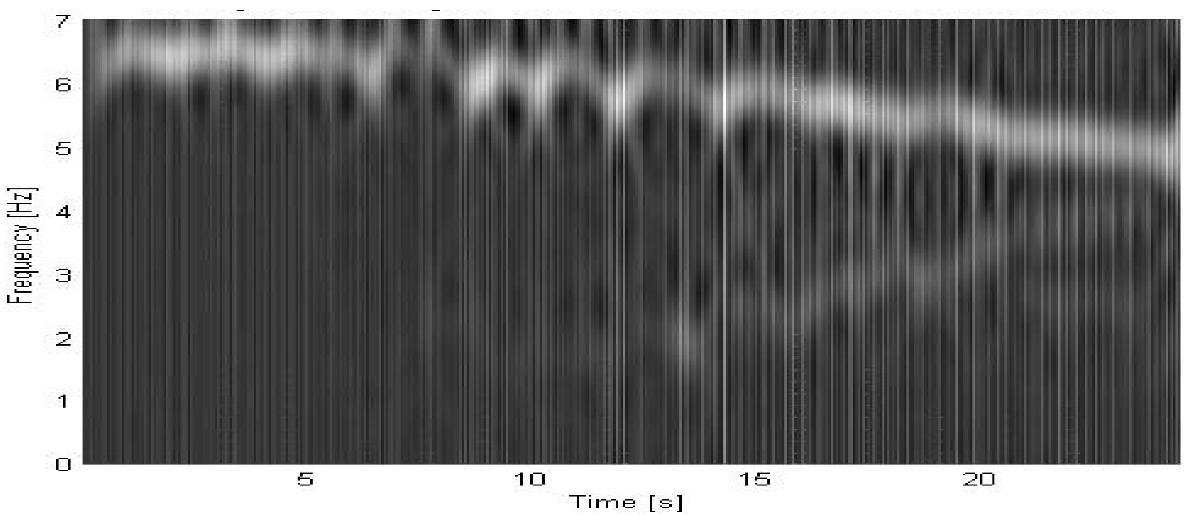

Figure 3.9: Choi Williams TFR for EEG seizure trace with a window of size 150 bins

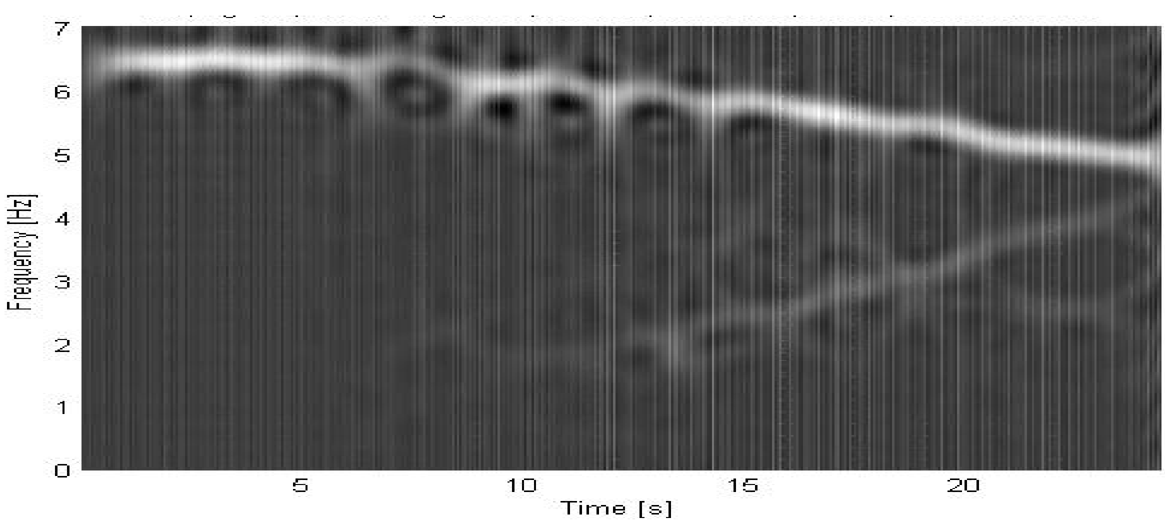

Figure 3.10:Choi Williams TFR for EEG seizure trace with a window of size 300 bins

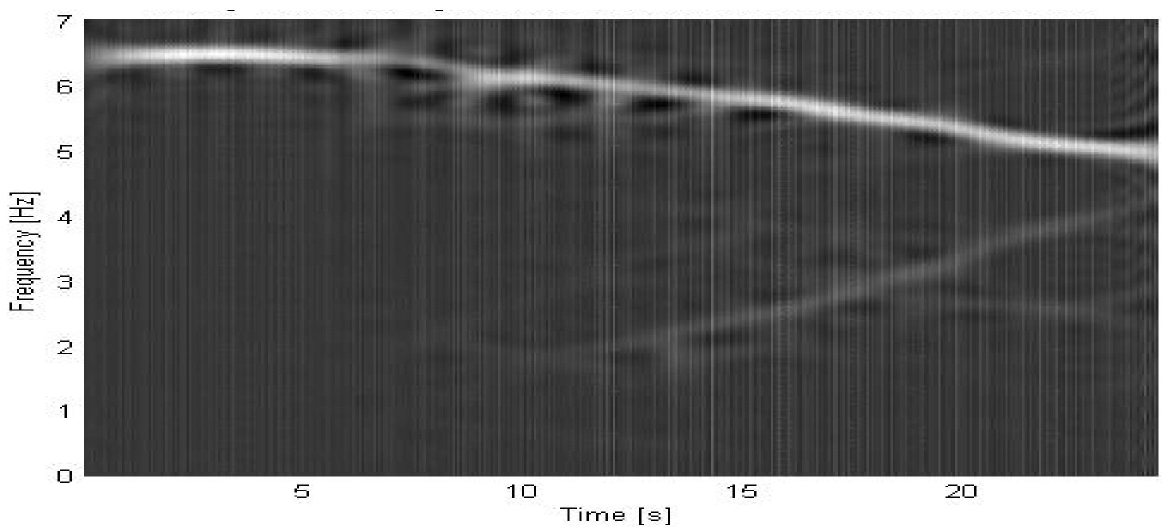

Figure 3.11:Choi Williams TFR for EEG seizure trace with a window of size 500 bins

From the figures it can be said that the Choi William representation with a window size of 500 gives a better representation when compared to other window sizes. The drawback of exponential kernel is that it can only reduce the cross terms close to the time and frequency center but for the cross term location on the ϑ and τ axis this kernel can do nothing. Also the parameter σ in the kernel function which is an important factor for improving resolution gives artifacts which are difficult to eliminate.

3.4.4 Zhao Atlas Marks Distribution (ZAM)

Zhao Atlaz Marks was proposed in 1990 by Y.Zhao, L.E.Atlas, and R.J.Marks to completely eliminate the effect of cross terms from the time frequency representation of signals [**53**].The ZAM time frequency distribution gives a good time and frequency domain resolution by reducing the cross terms to greater extent. It uses a cone shaped kernel and hence also called as cone shape distribution. The ZAM distribution uses the same TFR as the Choi William but with a cone shaped kernel function. The ZAM TFR with its kernel function is given by

$$ZAM(t,f) = \iint_{-\infty}^{\infty} A(\vartheta,\tau).\,\varphi(\vartheta,\tau).\,\exp\big(j2\pi(\vartheta t - \tau f)\big)\, d\vartheta d\tau \qquad (3.11)$$

Where $A(\vartheta,\tau)$ is the Ambiguity function and $\varphi(\vartheta,\tau)$ is the kernel function given by

$$\varphi(\vartheta,\tau) = \frac{\sin(\pi\eta\tau)}{\pi\eta\tau}\exp(-2\pi\alpha\tau^2) \qquad (3.12)$$

Where α is a adjustable parameter[**54**].

The advantage of this special kernel function is that it completely eliminates the cross terms. The ZAM time frequency representation with different number of frequency bins are shown in the figures 3.12 – 3.14. It can be seen from the figures that ZAM distribution with frequency bins size 500 is found to give good representation of seizure EEG trace.

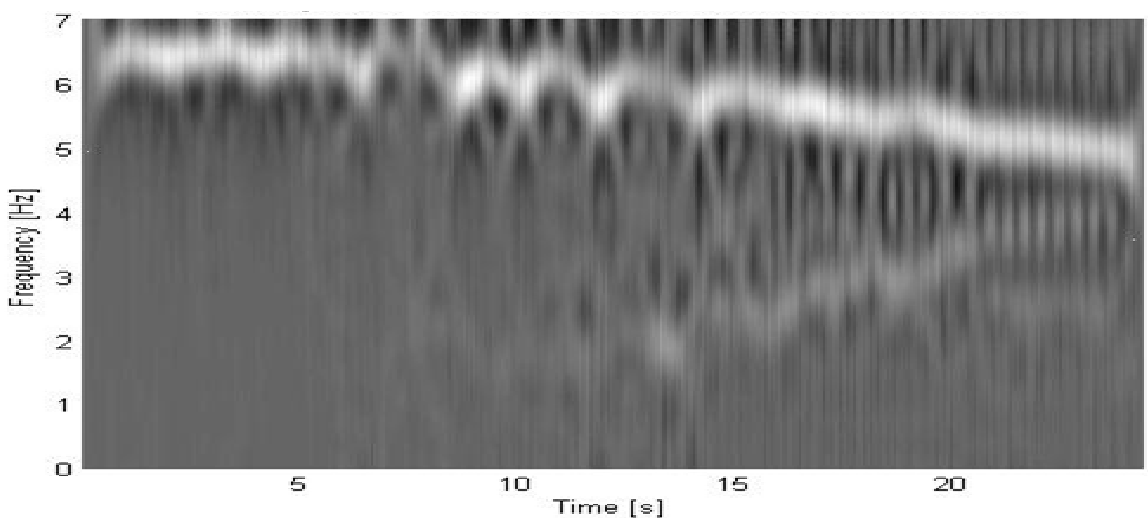

Figure 3. 12: ZAM TFR for EEG seizure trace with a window of size 150 bins

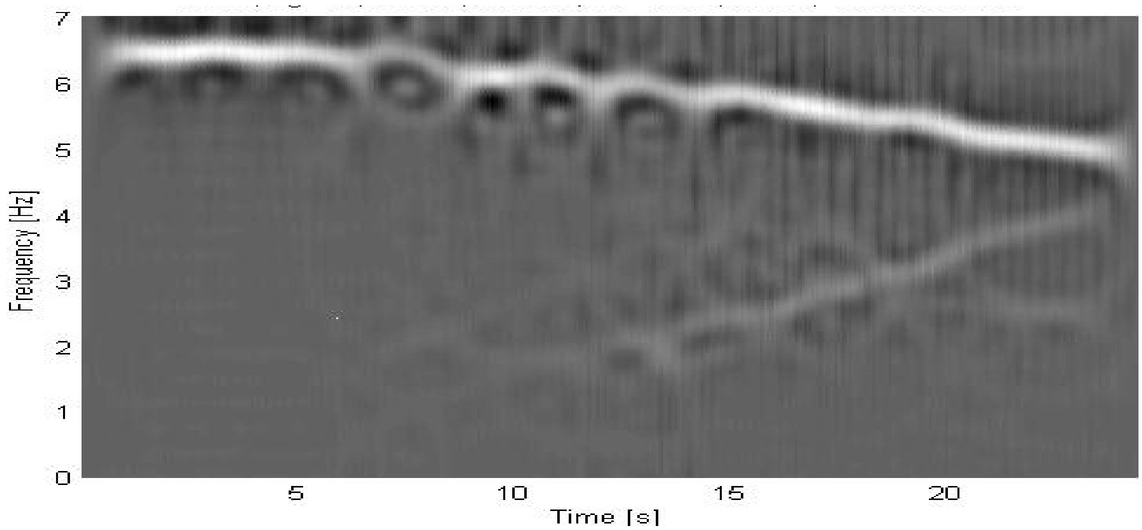

Figure 3. 13: ZAM TFR for EEG seizure trace with a window of size 300 bins

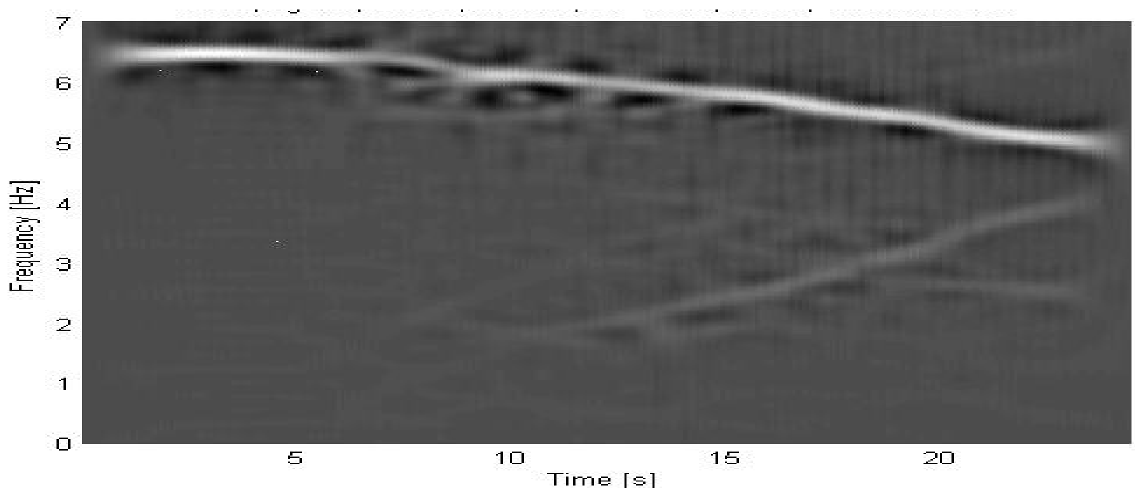

Figure 3.14: ZAM TFR for EEG seizure trace with a window of size 500 bins

3.4.5 Comparison and Conclusion

For comparison we have selected the best representation of seizure event by each Time frequency representation. It can be seen from the figures 3.15 – 3.18 that the STFT and Wigner Ville distribution give very poor representation of seizure trace. The Choi wiliams is found to give poor time resolution compared to ZAM. Also we can see several lines between 0-4 Hz in ZAM compared to all other TFR and hence we will be using ZAM distribution for our algorithm.

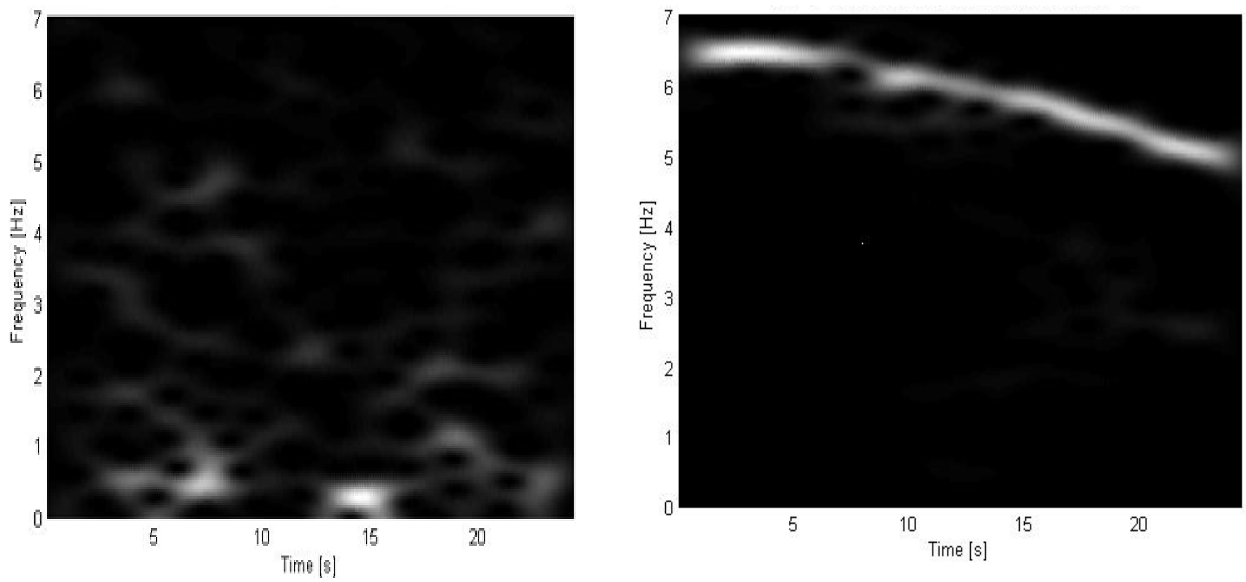

Figure 3. 15: STFT TFR for EEG non seizure trace (left) and seizure trace (right)

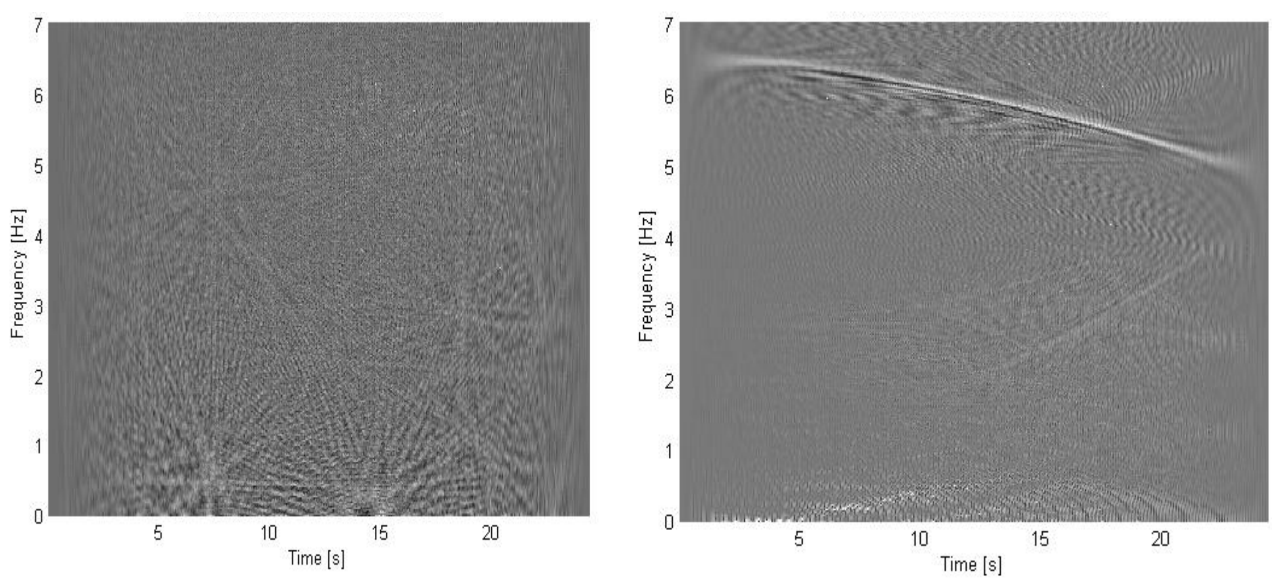

Figure 3. 16: Wigner Ville TFR for EEG non seizure trace (left) and seizure trace (right)

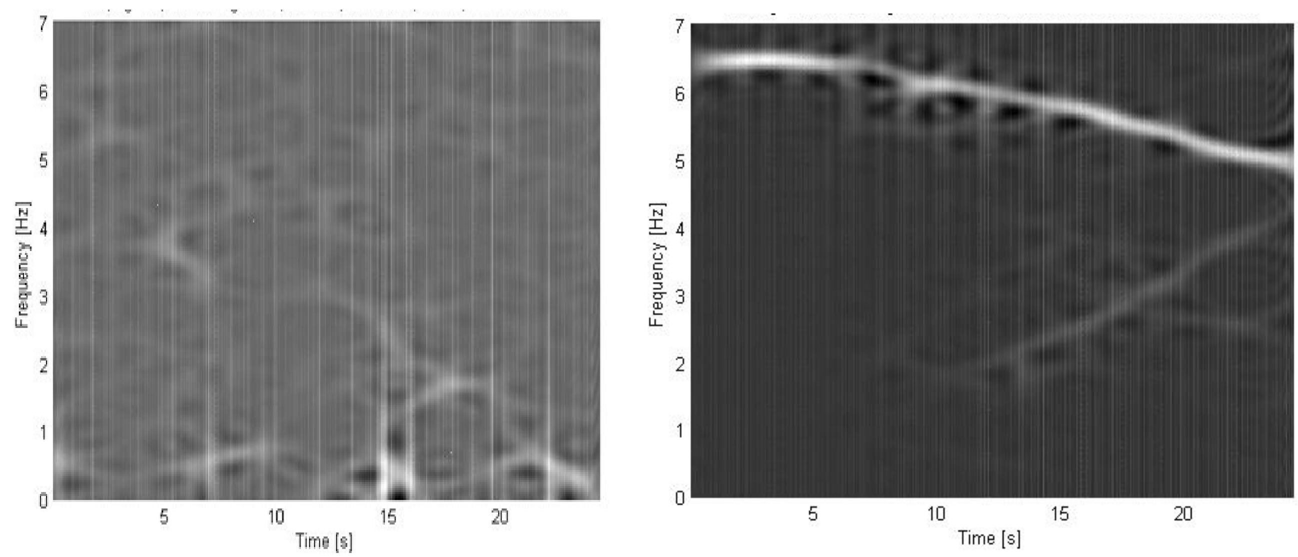

Figure 3. 17: Choi Williams TFR for EEG non seizure trace (left) and seizure trace (right)

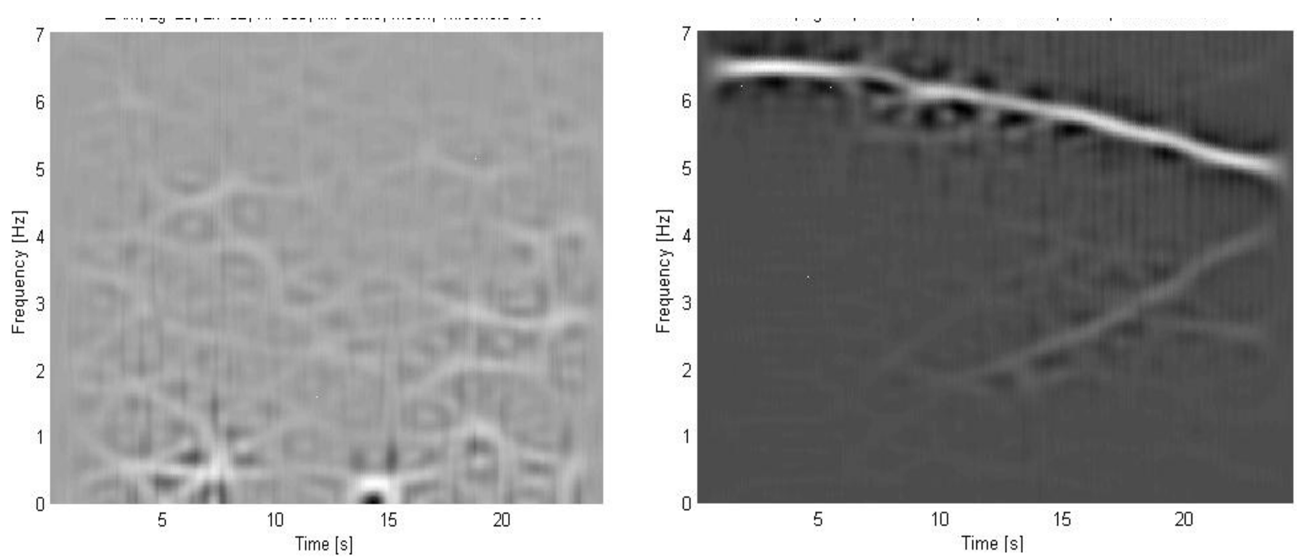

Figure 3. 18: ZAM TFR for EEG non seizure trace (left) and seizure trace (right)

Once the EEG trace is represented using ZAM TFR, we are going to perform Singular Value Decomposition on the TFR matrix to extract the signal information from the Time Frequency matrix.

3.5 Singular Value Decomposition

Singular Value Decomposition (SVD) is a popular factorization approach of rectangular real or complex matrices. The basic objective of SVD is to find a set of "typical" patterns that describe the largest amount of variance in a given dataset. In this thesis, we use the SVD decomposition on the time frequency distribution matrix **X** (MxN):

$$\mathbf{X} = \mathbf{U}\sum\mathbf{V}^{\mathbf{T}} \tag{3.13}$$

where $\boldsymbol{U}(M \times M)$ and $\boldsymbol{V}(N \times N)$ are orthonormal matrices, and $\boldsymbol{\Sigma}$ is an $M \times N$ diagonal matrix of singular values ($\sigma_{ij} \neq 0$ if $i = j$ and $\sigma_{11} \geq \sigma_{22} \geq \cdots \geq 0$). The columns of orthonormal matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are called the left and right Singular Vectors (SV), respectively. Note that matrices U and V are mutually orthogonal. The singular values (σ_{ii}) represent the importance of individual SVs in the composition of the matrix. The SVs corresponding to larger singular values provide more information about the structure of patterns contained in the data. As it can be seen from the figure 3.19 that the first Singular Value itself contains more than 60% energy of the signal. Hence we are using only the first Singular Vector corresponding to the first Singular Value as a feature vector for differentiating between the seizure and non seizure trace.

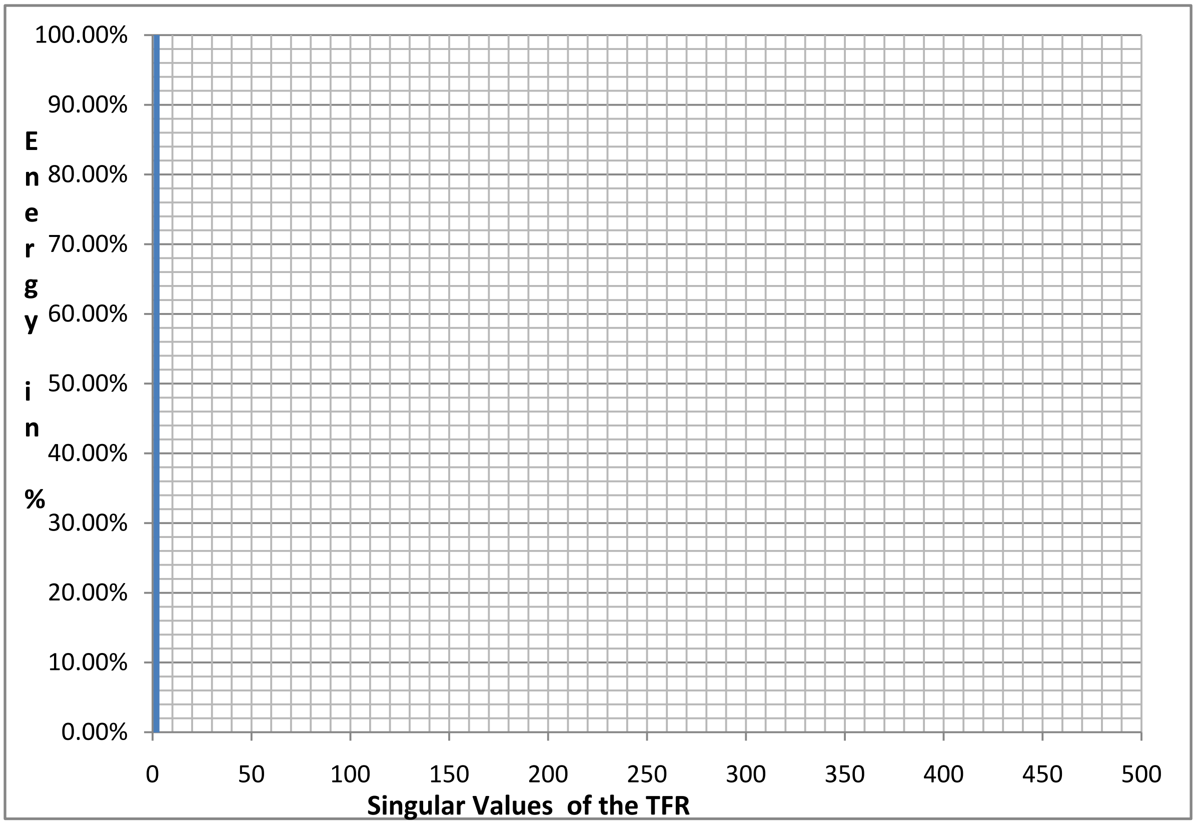

Figure 3. 19: Energy of the Singular values of TFR

3.6 Extracting Feature Vector

As we know that the singular values are orthonormal, which means that they have unit norm and hence their squared elements can be treated as probability mass functions (pmf) for different elements of the vector. For example the pmf of first columns of matrix U can be given as follows

$$F_u = \{u^2_{11}, u^2_{12}, \ldots\ldots\ldots\ldots\ldots, u^2_{1N}\} \quad (3.14)$$

From the above obtained pmf's we compute for histogram bins.

- The whole column data of the left singular vector is distributed in a non linear histogram bins. The reason for using non linear histogram bins is to focus more on the low frequency and high frequency information of the signal as the seizure events are related to an activity in the delta region (0-4Hz) . The histogram we are using in this research for the left singular vector has 17 bins which represent the frequency content of the signal. We have performed experiment with varying bins sizes and found 17 bins with non linear distribution of frequency information to be useful for classification purpose. The first 4 histogram bins represent information of frequency 0.5-1Hz, 1-2Hz, 2-3Hz and 3-4Hz. These histogram bins represent the characteristic vector to be fed to the linear discriminant network for discriminating a seizure event.

- In a similar way the column data for the right singular vector is distributed in histogram bins. But here we are using uniform bins as the right singular vector represents the information related to time and hence there is no point is distributing the data in a non linear way. In our research we are using 10 bins to represent the time information.

3.6.1 Left Singular Vectors as Feature Vectors

Previous researchers [**23**] have mentioned the use of both left and right singular vectors as characteristic features for discriminating between a seizure and non seizure event. In this research however we are using only Left singular vector for discriminating between different signals for the following reasons:

1. The right singular vector only shows the time information of the signal. It only shows the information at which time instant the seizure occurred. The seizure can occur at different time instant for different patient and even for the same patient may undergo seizure at different intervals of time.

2. It was also shown in that research [**23**] with an example of two signals which showed same left singular value plot for both the signals but showed different plots for right singular value and hence this is confirmed as a proof to establish that right singular value is necessary to discriminate between two different signals. However, the proof does not hold good when it comes to discriminating between a seizure and non seizure signal. This is because the difference in time singular value does not represent the seizure. Even though there appears to be difference between two signals in the example showed by the author, we say that both different signals belong to the same one group. The difference in the representation of time singular value only represents the time at which a seizure occurs. The seizure should be discriminated only on the basis of frequency.

To further strengthen our statement we present an example of a signal which represents the EEG of seizure undergoing patient. We get another signal from this seizure signal by time delaying it for 10 seconds.

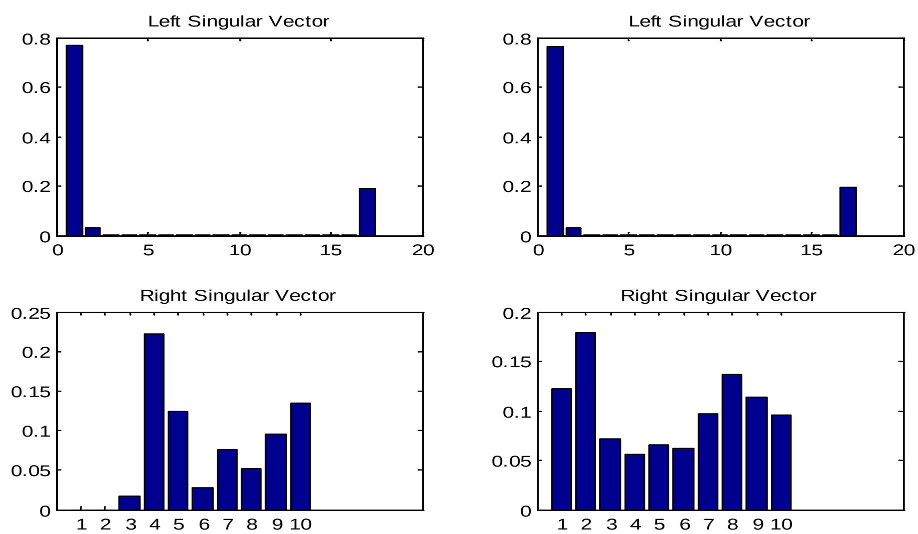

Figure 3.20: Histogram bins of EEG trace for seizure and its time shifted version

Both the signal undergoes the same steps for extracting the features. It can be seen from the figure 3.20 that the left singular value of both the signals remains the same but there is a change in the right singular value of the two signals. Thus the use of right singular value in discriminating the signals in detecting seizures is misleading and should be avoided.

3.6.2 Algorithm for Seizure Detection

To summarize the proposed algorithm for time frequency based seizure feature extraction comprises the following steps:

Step 1: Filtering

We are performing experiment on the low frequency signatures and any activity above 14Hz is filtered by passing the signal through a low pass filter with a cut off frequency of 14Hz.

Step 2: Down sampling

The data mentioned above is 23.6 seconds long and with a sample rate of 178.13Hz it has 4097 total number of samples. The sampling rate is reduced to reduce the computational load. The sampling rate here is reduced to 28Hz. Following the SyQuest rate this sampling rate is enough to analyze signals with frequencies less than 14Hz.

Step 3: time frequency representation

Zhao Atlas Marks (ZAM) distribution is used to represent the EEG signal in time frequency domain.

Step 4: Singular Value Decomposition

Applying singular value decomposition to the time frequency representation matrix and computing left and right singular values.

Step 5: Extracting Probability mass function

Since the columns of the matrix are orthonormal and hence the square of the elements can be considered as pmf’s .

Step 6: Histogram computing

From the probability mass function we compute histogram with 17 bins for the Left Singular Vector and 10 bins for the Right Singular Vector.

The figures 3.21 & 3.22 are for a seizure and non seizure trace corresponding to the first singular value. It can be seen from the figure that the Histogram corresponding to the Left Singular Vector easily discriminated between seizure and non seizure events. For a seizure trace it is found that the first and last bins of the histogram have large value and rest of the bins are almost empty, whereas for a non seizure trace the histogram bins are unevenly distributed.

If we consider the histogram bins for seizure and non seizure trace corresponding to the 2^{nd} singular value as shown in figures 3.23 & 3.24, it was found that even the Left Singular Vector for seizure trace is also unevenly distributed and hence the usage of other singular vectors reduces the overall detection accuracy. Hence, we are using the Histogram bins of the Left Singular Vector corresponding to the first singular value as the feature vector.

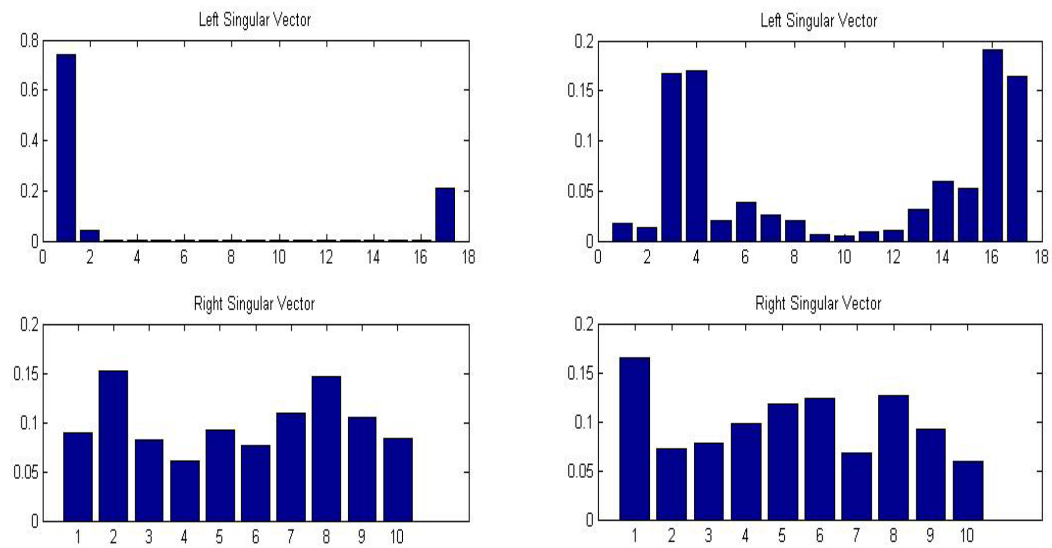

Figure 3. 21: (Sample 1) Pmf's of Left and Right singular vector corresponding to 1st singular value of a seizure (Left) and non seizure trace (Right)

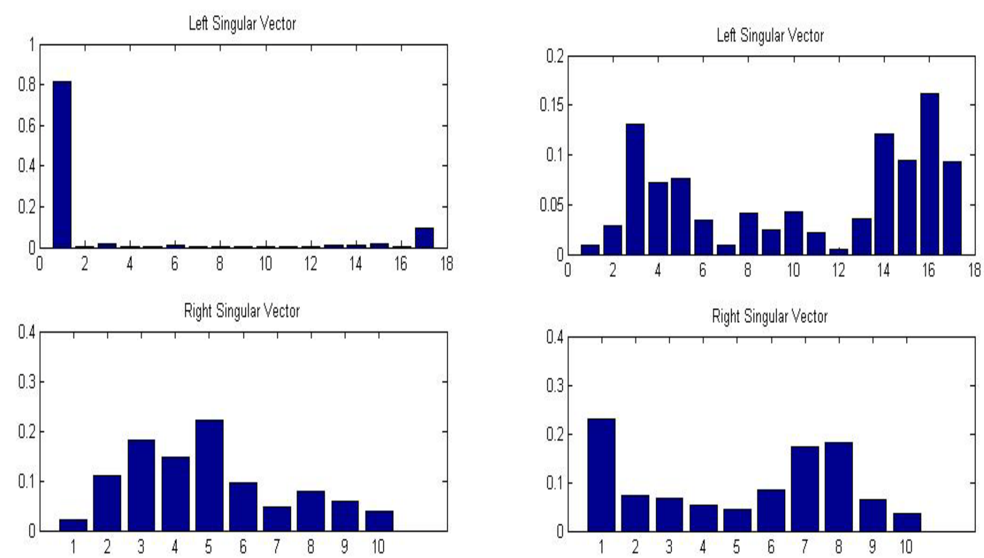

Figure 3. 22: (Sample 1) Pmf's of Left and Right singular vector corresponding to 1st singular value of a seizure (Left) and non seizure trace (Right)

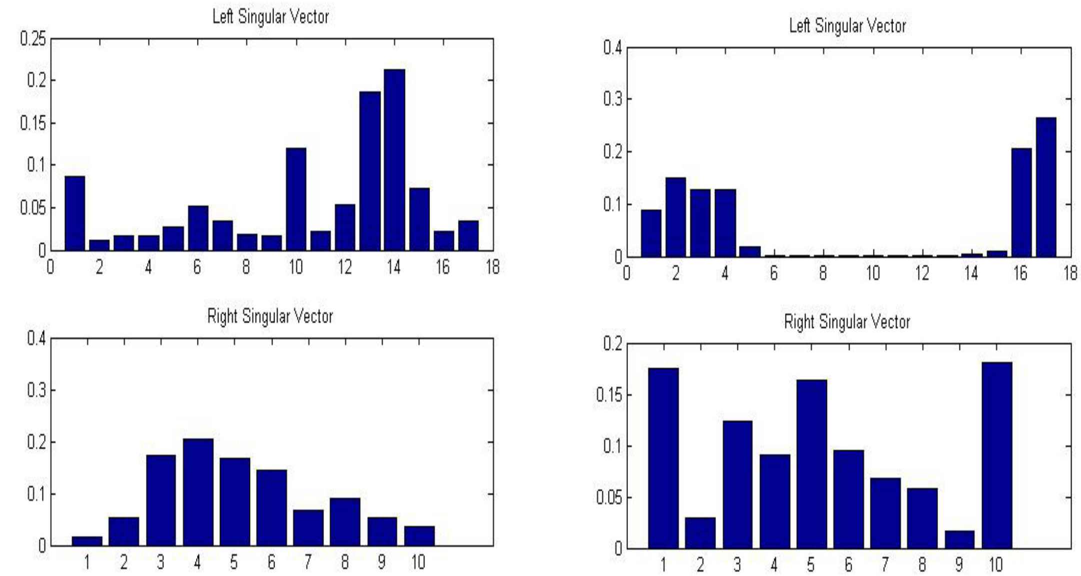

Figure 3. 23: (Sample 2) Pmf’s of Left and Right singular vector corresponding to 2^{nd} singular value of a seizure (Left) and non seizure trace (Right)

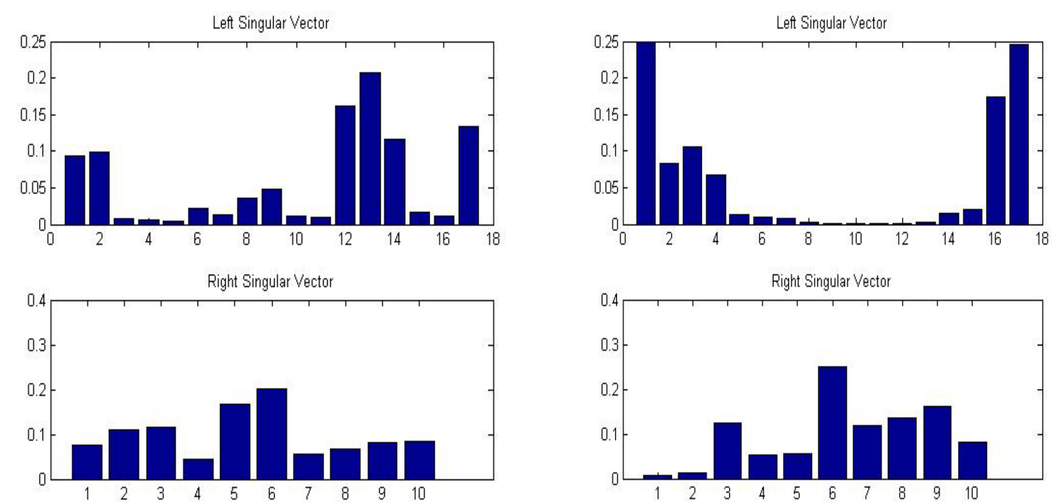

Figure 3.24: (Sample 2) Pmf’s of Left and Right singular vector corresponding to 2^{nd} singular value of a seizure (Left) and non seizure trace (Right)

The flow chart of the algorithm for EEG feature extraction is shown in the figure 3.25 below.

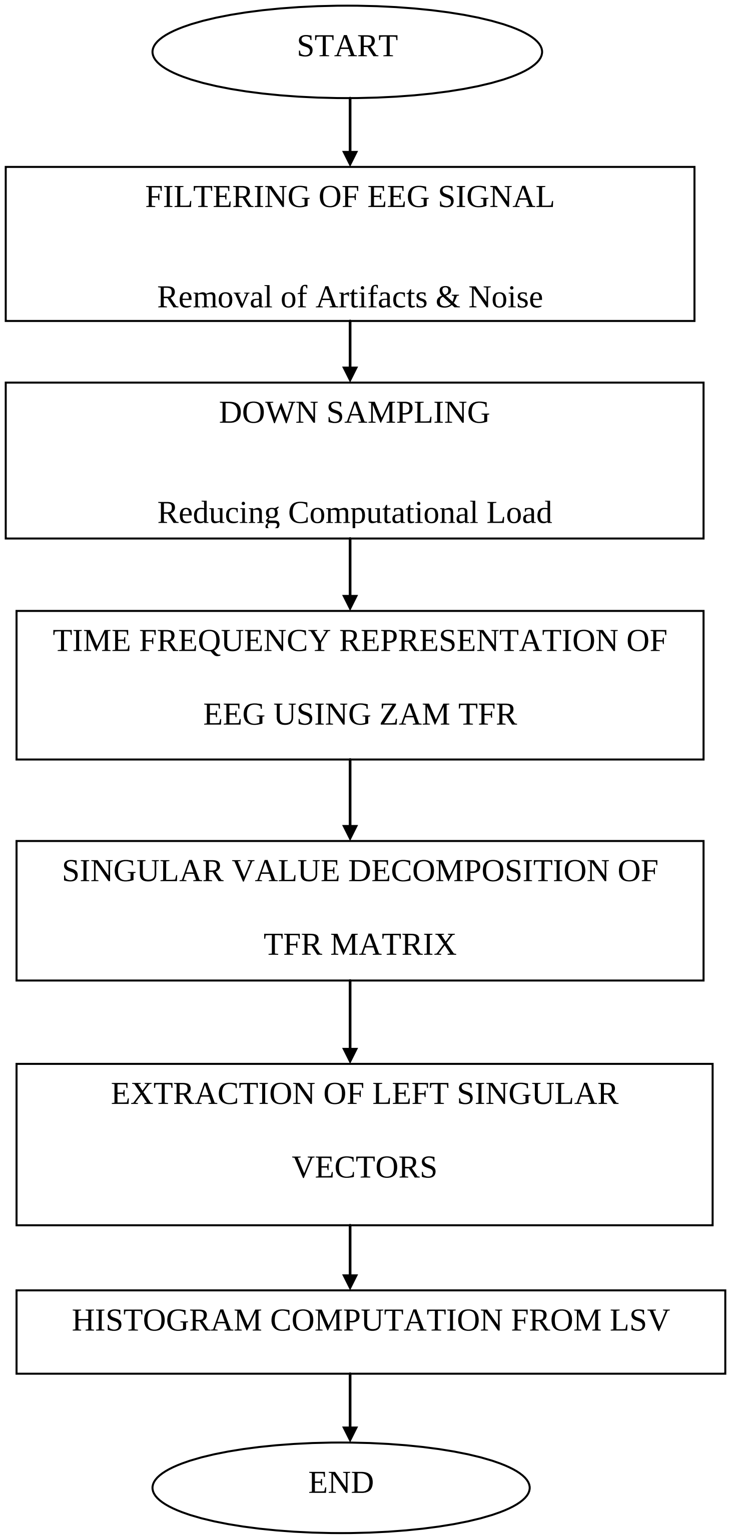

Figure 3. 25: Flow chart for feature extraction from EEG signal

3.7 Classification

After finding the features we now classify the EEG signals into seizure and non seizure traces. For this purpose we are using Linear Discriminant Analysis, which is very simple and effective technique for classifying the information in one of the two classes viz seizure and non seizure. It is found to be effective in pattern recognition case when the data set is large [**55**]. In contrast to Principal Component Analysis (PCA), which assumes each feature sample as a separate class the LDA assumes all the sample features belonging to the same group as a single class. The classification in LDA is then performed by minimizing the distance between the group and maximizing the distance among the groups and thus achieving maximum detection rates. Hence, PCA is found to be useful when dealing with small data sets only and for large data sets, as in our case LDA is best suitable for Classification [**55**].

3.7.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is one of the most commonly used dimension reduction technique. "LDA as classifier and as a feature extraction method has been used successfully in many applications including face recognition, other biometric techniques, finance, marketing, vibration analysis, etc"[**56**].

LDA was originally used for dimensionality reduction and works by projecting high-dimensional data onto a low dimensional space where the data achieves maximum class separability. The resulting features in LDA are linear combinations of the original features, where the coefficients are osbtained using a projection matrix **W**. The optimal projection or transformation is obtained by minimizing *within-class-distance (*between the signals of same group*)* and maximizing *between-class-distance* (between the signals belonging to different groups) simultaneously as shown in the figure 3.26, thus achieving maximum class discrimination. The optimal transformation is readily computed by solving a generalized eigenvalue problem.

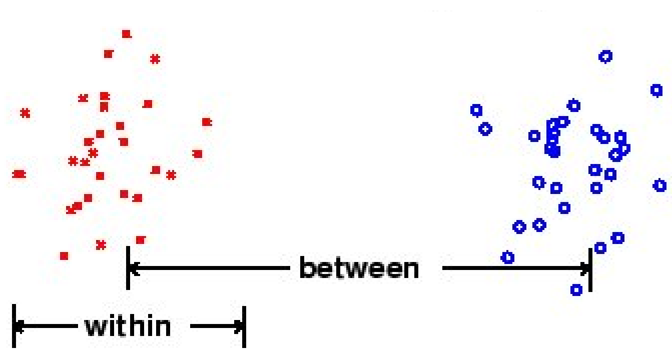

Figure 3. 26:Representation of Class separation in LDA

The initial LDA formulation, known as the Fisher Linear Discriminant Analysis (FLDA) was originally developed for binary classifications. The key idea in FLDA is to look for a direction that separates the class means well (when projected onto that direction) while achieving a small variance around these means. Discriminant Analysis is generally used to find a subspace with *M* - 1 dimensions for multi-class problems, where *M* is the number of classes in the training dataset.

More formally, for the available samples from the database, we define two measures: (i) *within-class* scatter matrix, given by:

$$S_w = \sum_{j=1}^{M} \sum_{i=1}^{N_i} (x_i^j - \mu_j)(x_i^j - \mu_j)^T \tag{3.15}$$

where $\boldsymbol{x}_i^j$ (dimension $\boldsymbol{nx1}$) is the i[th] sample vector of class j, $\boldsymbol{\mu}_j$ is the mean of class j, M is the number of classes, and N_i is the number of samples in class j.

The second measure (ii) is called between-class scatter matrix and is defined as:

$$S_b = \sum_{j=1}^{M} (\boldsymbol{\mu}_j - \boldsymbol{\mu})(\boldsymbol{\mu}_j - \boldsymbol{\mu})^T \tag{3.16}$$

where $\boldsymbol{\mu}$ is mean vector of all classes.

The goal is to find a transformation $\boldsymbol{W}$ that maximizes the between-class measure while minimizing the within-class measure. One way to do this is to maximize the ratio det($\mathbf{S_b}$)/det($\mathbf{S_w}$). The advantage of using this ratio is that if S_w is a non-singular matrix then this ratio is maximized when the column vectors of the projection matrix, $\boldsymbol{W}$, are the eigenvectors of $\boldsymbol{S_w}^{-1}.\boldsymbol{S_b}$ [**56**]. It should be noted that: (i) there are at most *M-1* nonzero generalized eigenvectors, and so an upper bound on reduced dimension is M-1, and (ii) we require at least *n* (size of original feature vectors) + M samples to guarantee that S_w does not become singular.

In the work discussed here, we use LDA to transform the PMF raw feature vector of dimension 17 (step 6 above) into a reduced feature (of projections) with a varying dimension between 1 and 17. We are using LDA here to classify the features obtained from the above algorithm into two different groups known as seizure and non seizure. The LDA algorithm at first assigns a group to a set of features belonging to the same class and when the algorithm is trained with the set of features available for training it classifies the test vector features to one of the group using Euclidean distance as a measure to to know to which group the given signal is closer to.

3.8 Experimental Results and Performance Comparision

From the available 200 traces, we used 45 traces from healthy individuals and 45 traces from subjects with seizures to train the LDA classifier. After estimating the LDA transformation matrix, we started the testing stage by projecting the test data over the LDA matrix, then using the Euclidian distances to classify a given test pattern as either a seizure or a non-seizure trace.

Out of the tested 110 samples, we were able to correctly classify 90% of traces. The experiment was carried again by randomly selecting different sets for testing and training. The recognition rates obtained for 10 trials were all very close to 90% (between 87% and 95%). For a given dataset, we show in Fig. 6 the changes in seizure detection accuracy as we vary the number of features used in the LDA analysis. We note that around 10 features are largely sufficient to represent the variations in the data.

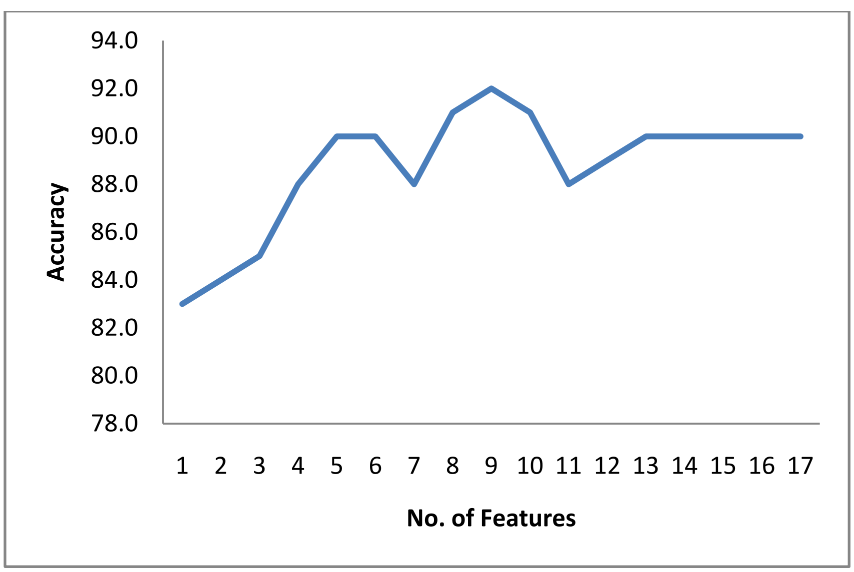

Figure 3. 27: Seizure detection accuracy as a function of the number of features from LDA

The Accuracy, sensitivity and specificity of a classifier are calculated as

Accuracy = $\frac{No.\ of\ Correct\ Detection}{Total\ No.of\ Traces\ of\ Healthy\ and\ Seizure\ events}$

Specificity = $\frac{No.of\ True\ Negatives}{No.\ of\ True\ Negative\ +No.of\ False\ Positives}$

Sensitivity = $\frac{No.of\ True\ Positives}{No.\ of\ True\ Positive\ +No.of\ False\ Negatives}$

The specificity of a classifier with 100% means that it identifies all healthy people as healthy whereas a sensitivity of 100% means that it identifies all sick people as sick. For our classifier we attained a specificity of 89.2% and sensitivity of 92.5%. The results achieved are comparable with the previous techniques.

The data used in the previous techniques mentioned in the table 3.1 is different from the data we have used in our research. Also, the detection accuracy in specified in terms of Good detection rate (GDR) and False detection rate (FDR). The GDR and FDR are given by

$$GDR = 100 \times \frac{GD}{R}\ and\ FDR = 100 \times \frac{FD}{GD + FD}$$

Where GD and FD are total number of good detection and false detection respectively and R is the total number of seizures correctly recognized by the neurologist. It can be seen that the detection accuracy here is dependent on the accuracy of the neurologist in predicting seizure from the raw EEG data. It was found in a research published by Clinical Neurology that the expert neurologist reports in the past were found to be 94% accurate[**57**]. Based on this accuracy of the neurologist we have converted the GDR and

FDR mentioned in the previous papers to sensitivity and specificity measures. We present in Table 3.1 a summary of the results we obtained showing that our proposed approach outperforms previously discussed techniques.

Technique used for seizure detection	Detection Accuracy	Sensitivity	Specificity
Auto Correlation technique proposed by A. Lieu	54%		
Basic Spectral technique proposed by J. Gotman	42%		
SSA technique proposed by P. Celka	85%		
DFSV technique proposed by H. Hassanpour	86%		
Back propagation neural network trained features by Ardalan Aarabi	79.7%	74.1%	70.1%
Our proposed technique	90%	92.5%	89.2%

Table 3. 1: Performance Comparison

3.9 SECTION SUMMARY

In this section we have discussed a time frequency based seizure detection technique which uses the EEG signal and extracts the left singular values from the time frequency matrix of the EEG signal to train the LDA. The different types of time frequency representation of EEG signal are discussed and Wigner ville distribution is selected to represent the EEG signal in time frequency domain as it is giving sharp features related to seizure trace of EEG signal. The result of the TF-LDA algorithm gives an average accuracy of 90% with sensitivity and specificity of 92.5% and 89.2% respectively. In the next chapter we are going to discuss about the detection of seizures based on Electrocardiogram (ECG) signals.

CHAPTER 4

SEIZURE DETECTION BASED ON ECG SIGNAL

4.1 Introduction

In recent years a number of algorithms for the detection of seizures based on electroencephalogram (EEG) have been proposed. More importantly, recent work has shown that in a number of cases, seizures are often associated with changes in heart and respiration rate[**58**]. The affect of complex seizures can be found in the cardiovascular system hence, seizures may also appear as variations in the cardiac rhythm[**58**]. In particular Seizures commonly may produce asystole, sinus bradycardia, and other disturbances in the normal ECG rhythm[**59**]. Even though, there exists an extended body of work in the seizure detection based on EEG, much less work can be found in the detection of seizures using ECG traces.

In this thesis, we propose to combine the information from both EEG and ECG in the robust detection of seizures. Before describing our proposed algorithm for detection of seizures based on ECG signals, we will first start by explaining effect of seizures on the heart.

4.2 Anatomy of the Heart

To get a good insight and understanding of ECG, we will first explain the basic anatomy of the heart.

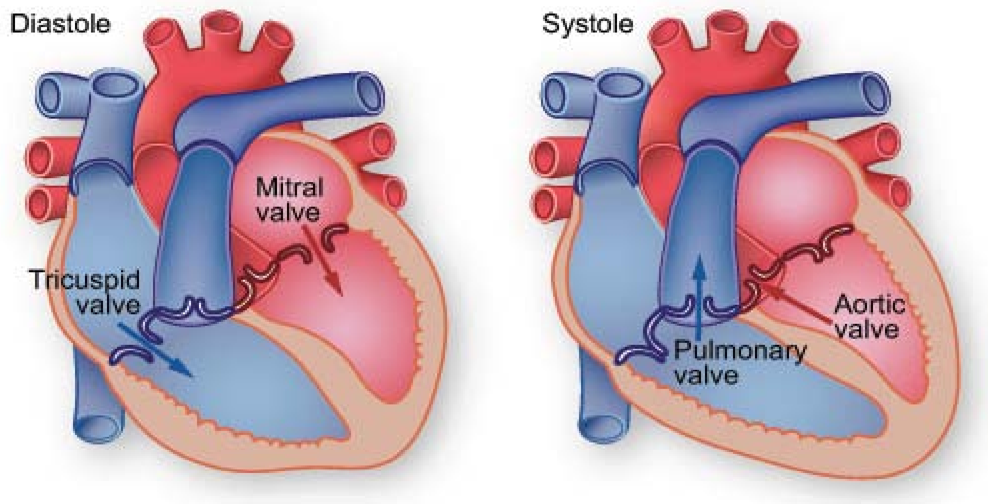

Figure 4. 1: Heart Valves [**60**]

The Heart is a 4 chambered muscle whose function is to pump blood throughout the body[**61**][**62**]. The upper chambers are called the left and right atria and the lower chambers are called left and right ventricles. A wall of muscle called septum separates the atrias from the ventricles. Together there are four valves which regulates the flow of blood through the heart. These are:

- The Tricuspid valve :

  This valve regulates the flow of blood between the right atrium and the right ventricle. The blood entering through this valve is deoxygenated blood received from the body into the right atria. This blood is then pushed into right ventricle through the valve.

- The Pulmonary valve :

  This valve channels blood from the right ventricle into the pulmonary arteries which carry the de oxygenated blood into the lungs for oxygenation.

- The Mitral valve :

The oxygenated blood from the lungs enters the left atrium and passes to the left ventricle through this valve.

- The Aortic valve :

  The oxygenated blood from the left ventricle is pumped throughout the body by passing it into the Aorta which is seen as the largest artery in the human body [**60**].

4.3 Measurement of Electrical Activity Using ECG

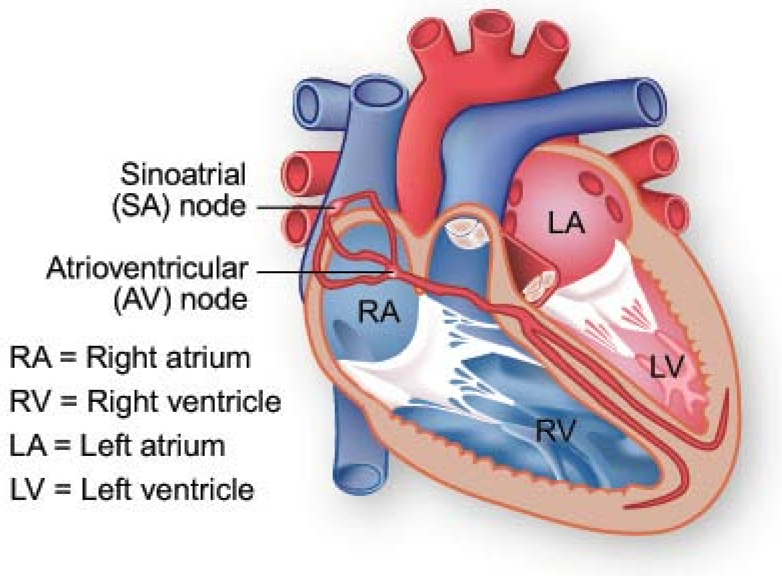

Figure 4. 2: Heart Valves [**60**]

The Electro Cardiogram (ECG or EKG) is a widely used diagnostic tool for measuring the electrical activity of the heart. It records the electrical activity of the muscles which causes the pumping of the heart and depicts it as a series of graph like tracings or waves. ECG traces help in monitoring the functioning of heart and reveal important information about any abnormalities that may exist.

The ECG represents the electrical activity of the heart that results due to the motion of the cardiac muscle myocardium which causes the heart to contract. In [**60**], the author states that the network of nerve fibers coordinates the contraction and relaxation of the cardiac muscle tissue to obtain an efficient, wave like pumping action of heart[**60**]. This contraction and relaxation of cardiac muscle is carried our throughout the lifetime of a human being and as a result blood flows through the heart and the process of oxygenation of blood is carried out.

The physiology of the heart together with respect to the contraction and relaxation of the muscles with some key elements is shown in the figure 4.2. The Sinuatrial node (SA) is known as the natural pacemaker of the heart. The SA node triggers an electrical impulse which results in a heart beat. This impulse thus passes through the atria resulting in contraction of atrium muscles and reaches the Atrioventricular node (AV) which triggers another pulse causing the ventricle muscles to contract.[**63**]

The trigger from the AV node is then received by the bundle of His which divides the triggering pulse between the right and left ventricles resulting in contraction and relaxation of right and left ventricles. This series of waves causing contraction and relaxation produces a wave like rhythm and this rhythm can be recorded through different tools available.

These electrical signals are recorded by placing electrodes on top of the body strategically to detect the electrical activity produced by the heart. The ECG waveform obtained is shown in figure 4.3.

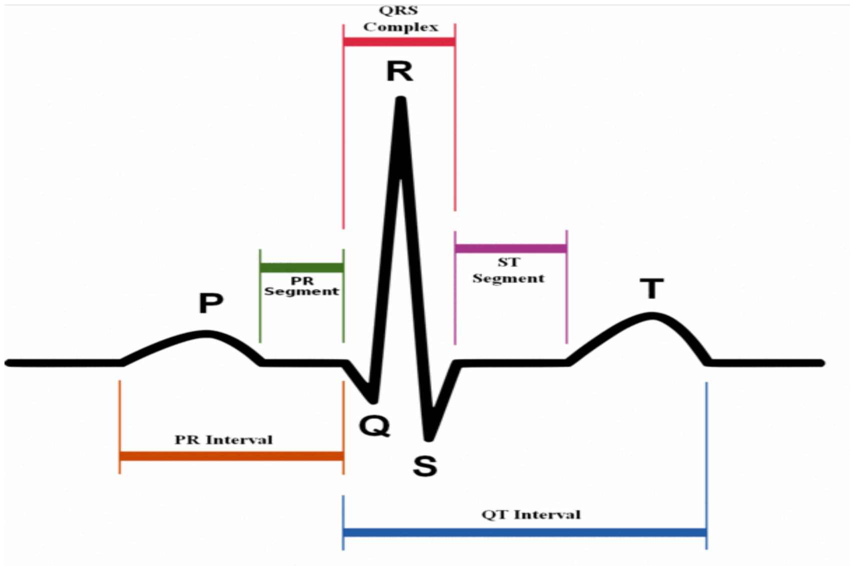

Figure 4. 3: ECG waveform [**64**]

The normal ECG begins with a P-wave which indicates the discharge of the sinoatrial node (SA). It represents the depolarization of the atria. The normal amplitude of the P-wave should not exceed 0.25 mV and duration of 0.11 sec[**65**].

The period of time from the onset of P-wave to the onset of Q-wave is called as PR interval. “It indicates the time between the onset of atrial depolarization and the onset of ventricular depolarization. The normal range of the PR interval lies between 0.12 and 0.20 sec.” [**66**].

“The QRS complex represents the ventricular depolarization. The duration of QRS complex lies between 0.06 and 0.1 sec. This short duration indicates that ventricular depolarization normally occurs rapidly”[**66**].

“The QT interval represents both the time for ventricular depolarization and repolarization to occur. It can range between 0.2 and 0.4 sec” [**66**].

4.4 Effects of Seizures on ECG Pattern

Seizures produce various effects on the cardiovascular function of the heart. These directly influence the central autonomic network thus controlling the heart rate and rhythm. It was shown that the patients affected with seizures have increased heart rate and several changes in the ECG rhythm. These changes are discussed below:

- Effect on the RR interval:

  A seizure often causes decrease in the RR interval. In the research discussed in [**67**], the author mentions that of the 24 patients evaluated, 92% of seizures were associated with an increased heart rate. It was also found in a recent study of 145 seizure events that seizures associated with onset tachycardia (increase in heart rate) occurred in 86.9% of all seizures, whereas bradycardia(decrease in heart rate) was documented in 1.4% and the remaining 11.7% seizures showed no change in the heart rate.[**68**].

- Effect on the PR interval:

  The PR duration is also effected during seizures as discussed by Stephen Oppenheimer[**69**]. A case has also been reported in [**70**] where patients effected with seizures were reported to have an increase in the PR duration.

- Effect on the P height:

  In [**69**] , the author states that changes in heart rate of the seizure affected patients are also accompanied by changes in p wave morphology[**69**].

- QRS interval

  The QRS interval is found to be unchanged during seizure interval[**71**].

- Effect on the QT interval

The QRS intervals were found to be unaffected by seizures [**71**]. A longer QT interval was reported in patients affected with seizure. In particular SUDEP (Sudden Unexpected Death in Epilepsy) is associated with longer QT interval. The QT interval has been used as an efficient feature for prevention of SUDEP[**72**]. In simple terms very long QT intervals leads ultimately to the person's death [**73**].

4.5 ECG database

The database used in the research is available on MIT database (http://physionet.org/physiobank/database/). The report on seizure was based on the analysis of data from 11 partial seizures recorded in patients ranging from 31 to 48 years old[**74**]. The non seizure database includes 18 long term ECG recordings of patients ranging from 20 to 50 years. The sampling rate of the data is 200Hz. A sample of original ECG signal is shown in figure 4.4.

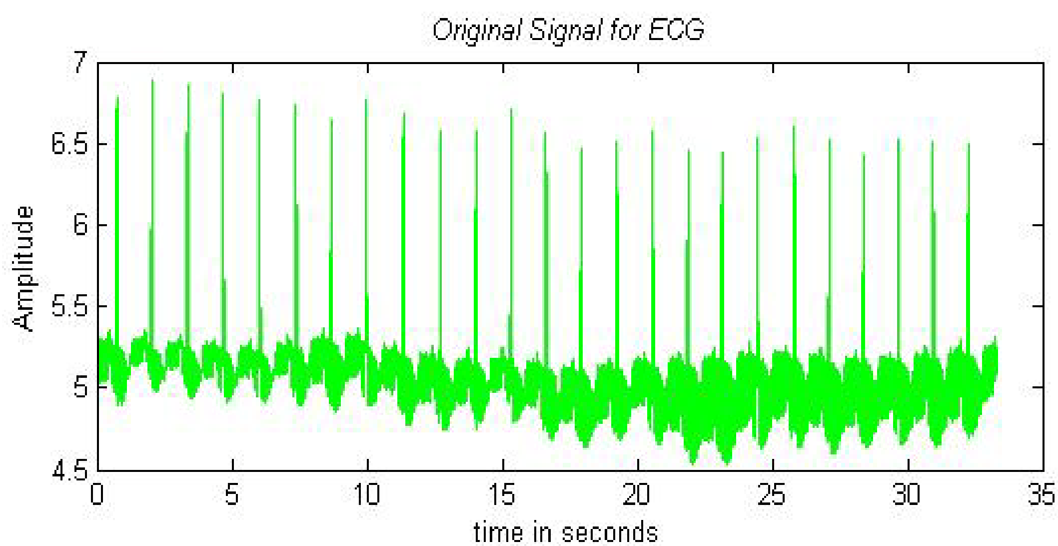

Figure 4.4: Original ECG signal

4.6 Extraction of Features from ECG Signals

Previous work on seizure detection has focused mainly on using RR intervals. In most studies, the different factors discussed above have not been used to their full extent in developing robust seizure detection algorithm. In this research, we focus on a whole set of features that were shown to be closely related to seizure occurrence. We then use these features to train and classify the ECG data using simple linear discrimination analysis. For our study above and the different discussions made with the KFUPM clinic here, we decided to use the following features:

1) R-R interval mean
2) R-R interval variance
3) P height mean
4) P-R duration
5) Q-T duration

These 5 features were found to be very effective in discriminating an ECG signals containing seizure and non seizure traces.

4.6.1 Wavelet Decomposition of ECG Signal:

To extract the R-R interval from the ECG signal as well as the other P,Q,S,T waves, we decompose the given ECG signal using the traditional wavelet transform.

The Wavelet transform has been used very frequently in different signal processing applications. The Wavelet Transform plays a crucial role in signal analysis as it is usually used to find hidden frequency content in a given signal which is not otherwise visible directly from time domain representation. Wavelet analysis consists of decomposing a signal or an image into a hierarchical set of approximations and details. The levels in the hierarchy often correspond to those in a dyadic scale. From the signal analyst's point of view, wavelet analysis is a decomposition of the signal on a family of analyzing signals, which is usually an orthogonal function method. From an algorithmic point of view, wavelet analysis offers a harmonious compromise between decomposition and smoothing techniques[**75**]. The wavelet analysis is performed in a similar way to the STFT, in the sense that the signal is multiplied with a function, similar to the window function in the STFT, and the transform is computed separately for different segments of the time domain signal. However, there are two main differences between the STFT and the CWT[**76**].

- "The Fourier transforms of the windowed signals are not taken, and therefore single peak will be seen corresponding to a sinusoid, i.e., negative frequencies are not computed"[**76**].
- "The width of the window is changed as the transform is computed for every single spectral component, which is probably the most significant characteristic of the wavelet transform"[**76**].

The continuous wavelet transform of given signal x(t) is given by

$$X(a,b) = \frac{1}{\sqrt{a}} \int x(t) \, . \, \psi\left(\frac{t-b}{a}\right) dt \tag{4.1}$$

Where a and b are dilation of the wavelet and time translation respectively. It can be thus understood from the equation that the wavelet transform of a signal decomposes the signal and gives collection of shifted and stretched versions at different scales.

In order for the estimation of ECG parameters from the ECG signal a proper selection of the wavelet is required. This choice leads us to the use of Biorthogonal wavelet as it satisfies the properties mentioned in [**77**] which suggest “the basis function to be symmetric/antisymmetric. A symmetric basis will enable the detection of peak of wave as an extrema. In case of antisymmetric basis, the peak of the wave is detected as a zero crossing. Also, it is desirable that the basis have a minimum number of sign changes which will simplify the steps in the parameters estimation algorithm” [**77**].

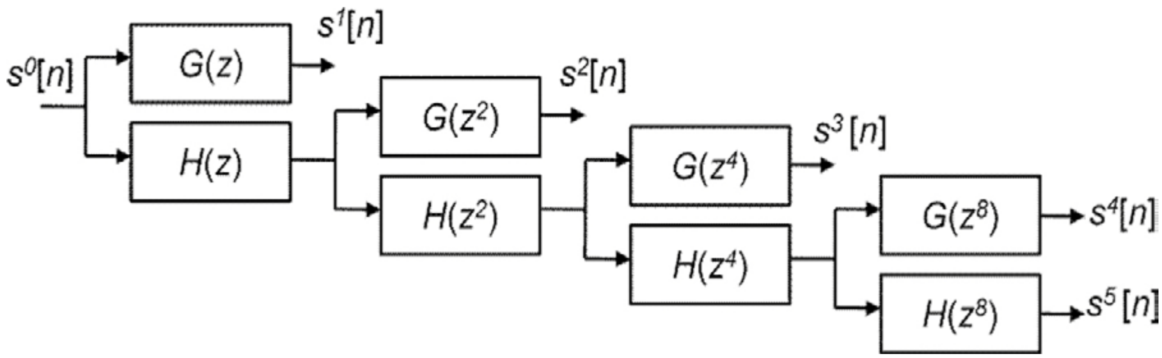

Figure 4. 5: Wavelet Decomposition tree for ECG signal

The ECG parameters are derived by the wavelet decomposition tree. At each stage the signal is decomposed into approximate (low pass) and detailed (high pass) coefficients. The low pass output of the signal is further decomposed into low pass and high pass. The process of decomposition is repeated for 4 time and when an ECG signal is passed through each of the wavelet filters whose scales range from 2^1 to 2^4, as shown in figure 4.5. The detailed and approximate signals are obtained. The different type of biorthogonal wavelets available in MATLAB are shown in figure 4.6 .The type of wavelet we are using in our research is bio 2.4 as it closely resembles the ECG signal.

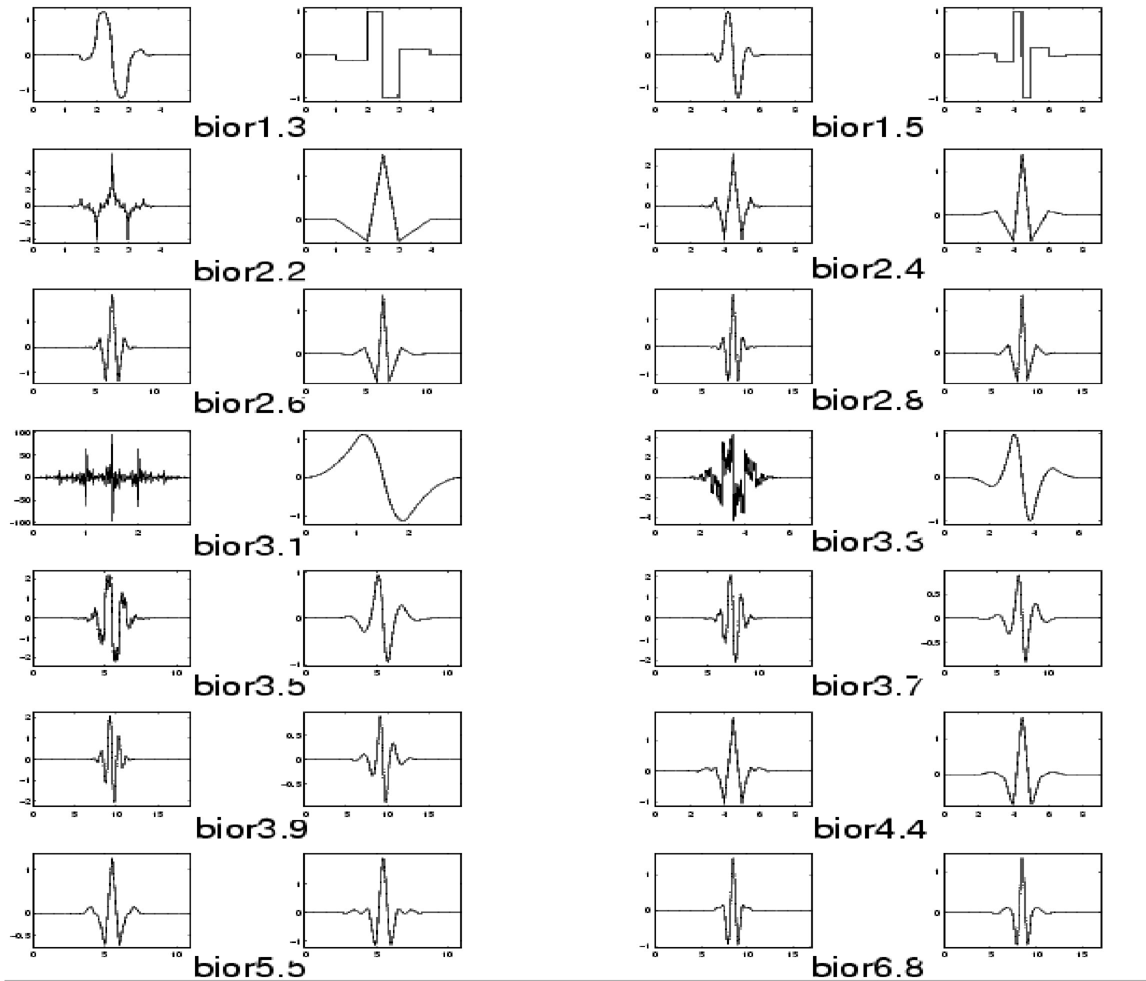

Figure 4. 6: Types of Biorthogonal wavelets in MATLAB [**75**]

"Wavelet transformation has shown to be substantially noise proof in ECG segmentation and thus appropriate for ST-T segment extraction. The signal was decomposed into 4 scales ranging from 2^1 to 2^4 . It was found that the wavelet transform at small scales reflects the high frequency components of the signal and, at large scales, the low frequency components. The energy contained at certain scales depend on the center frequency of the used wavelet"[**63**].

"The 2^4 scale of the wavelet transformed ECG signal is used to detect the R-peak because most energies of a typical QRS-comples are at scales 2^3 and 2^4. "The high frequency noises like the electric line interference, muscle activity, bowel movement activity, electromagnetic interference is concentrated in the lower scales of 2^1 and 2^2, while the levels 2^3 and 2^4 constitute for less noise compared to the lower scales. Thus it was summarized in that the frequency of QRS complex is mainly present in the 2^3 and 2^4 scales"[**63**]. As the 2^4 scale is found to have less noise compared to 2^3 , which can also be seen from the figure , we choose 2^4 scale for extracting R peaks in our project. The wavelet decomposed ECG signal is shown in figure 4.7

We then extract the R peaks from the 2^4 scale by setting some threshold using Tompkings method[**78**]. Once the R peaks are extracted we then extract the PQST peaks from the ECG wave using the Tompkins method which will be discussed in the following section.

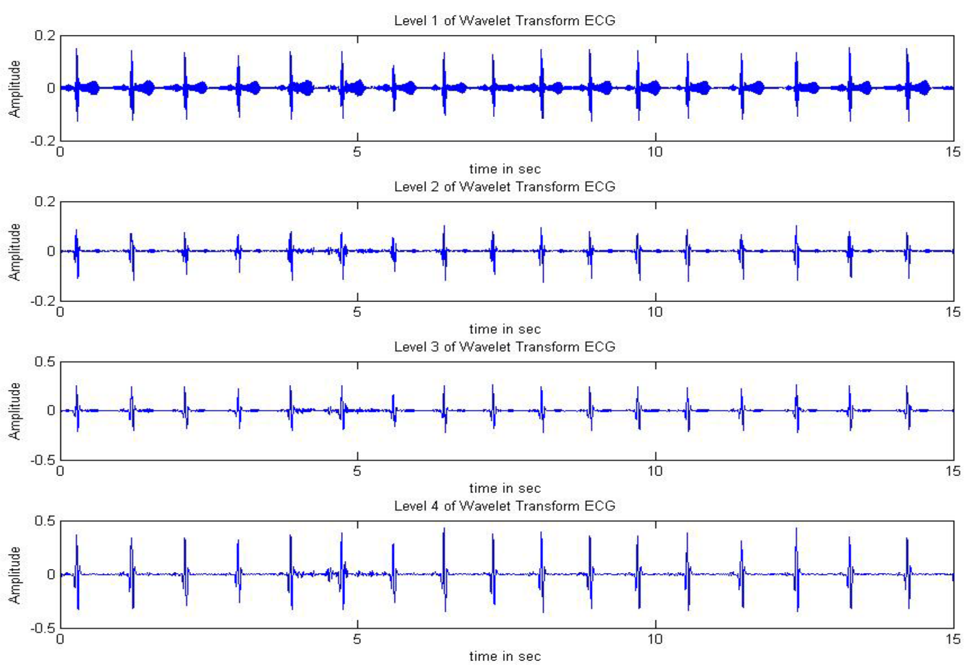

Figure 4. 7 Wavelet transformed ECG signal at different levels

4.6.2 Feature Extraction Algorithm:

Step 1: (ECG Signal Filtering)

The ECG data of length 60 seconds is used for analysis. This length of ECG data was found to be adequate in the previous research work[**34**]. The original ECG signal is shown in the figure 4.8. The data consists of many artifacts and noise due to the presence of power line interference, bowel movements also called EGG movement, muscle activity that gets captures along with the measured ECG signal , Electromagnetic interference. So in order to remove this noise we have to pre process the ECG signal before using it for further processing. This is done by using a simple FIR filter.

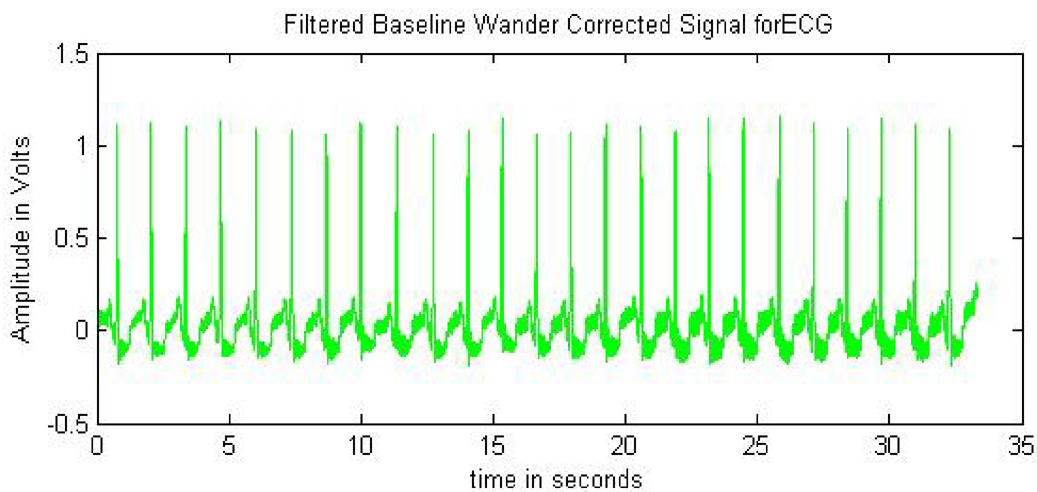

Figure 4. 8: Filtered and Baseline wander corrected ECG signal

Step 2: (Baseline Wander Correction)

Baseline wandering is also considered as an artifact which affects the measuring of ECG parameters. The respiration, electrode impedance change due to perspiration and increased body movements in most of the ECG are the main causes of the baseline wandering. In order to remove baseline wandering we pass the filtered signal through a median filter of length 200ms that remove the QRS complexes. The filtered signal is again passed through a median filter of length 600ms to remove the T wave. The filtered signal obtained in step 2 is then subtracted from the filtered signal obtained in step 1 which gives us the baseline wander eliminated signal. The filtered and baseline wander corrected signal is shown in figure 4.9.

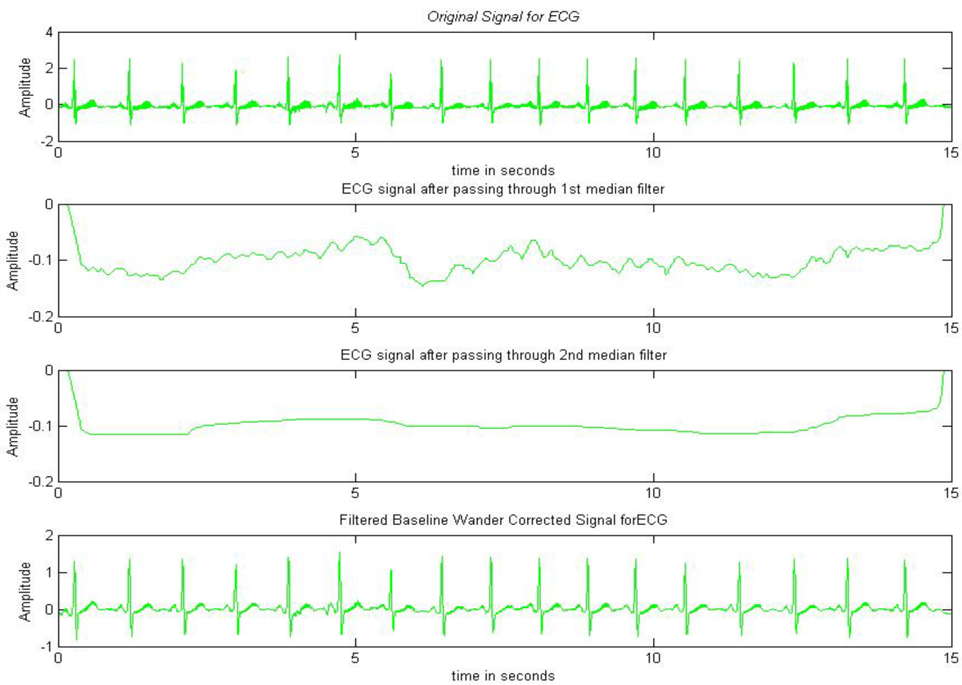

Figure 4. 9: Different steps in filtering ECG signal

Step 3: R peak detection

After getting the corrected ECG signal from step 2, R-peak detection algorithm is applied on the ECG signal. The detection of R-peak is based on threshold level to calculate maximum amplitude in the ECG waveform. The R-peak detection was done in the time scale domain at level 2^4. Same level is used to detect other key points in the ECG waveform.

Step 4: PQST detection

The PQST waves are then detected using the Tompkins method[**78**]. “After detecting R-peak, the first inflection points to the left and right are estimated as Q and S respectively. After estimating the S-point, J-point was estimated to be the first inflection point after S-point to the right of R-peak. T-peak was estimated to between R-peak+400ms to J-point +80ms. Similarly K-point was estimated to be the first inflection point after Q on the left side of the R-peak, and P-point was estimated to be the first inflection point after K-point on the P-peak side"[**63**].

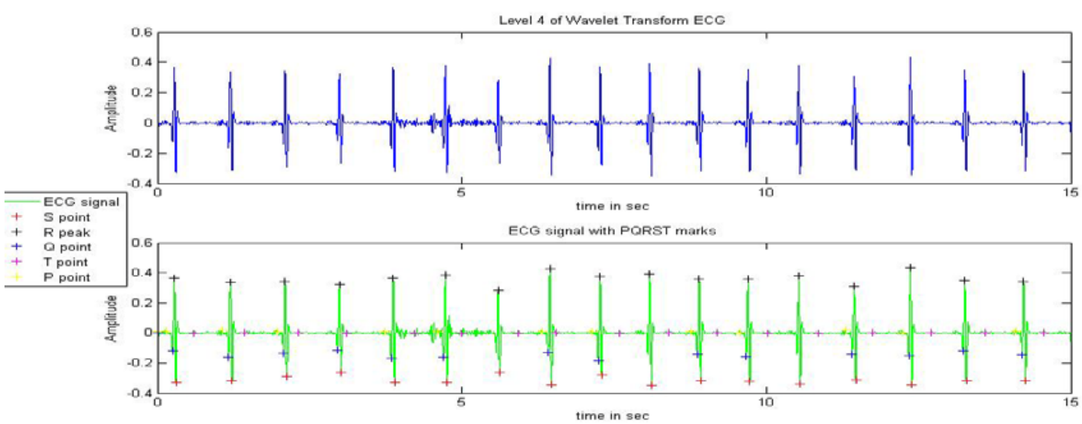

Figure 4. 10 Detected PQRST peaks from the ECG signal

Step 5: Feature Extraction

After getting all the required waves of ECG we now calculate the different features required for classification of ECG signals. We extract the RR-mean, RR-variance, P peak mean, QT duration mean, PR duration mean.

4.7 Flow Chart of Seizure Detection Algorithm

The Flow chart of the above mentioned seizure detection algorithm is shown in the figure 4.8 below.

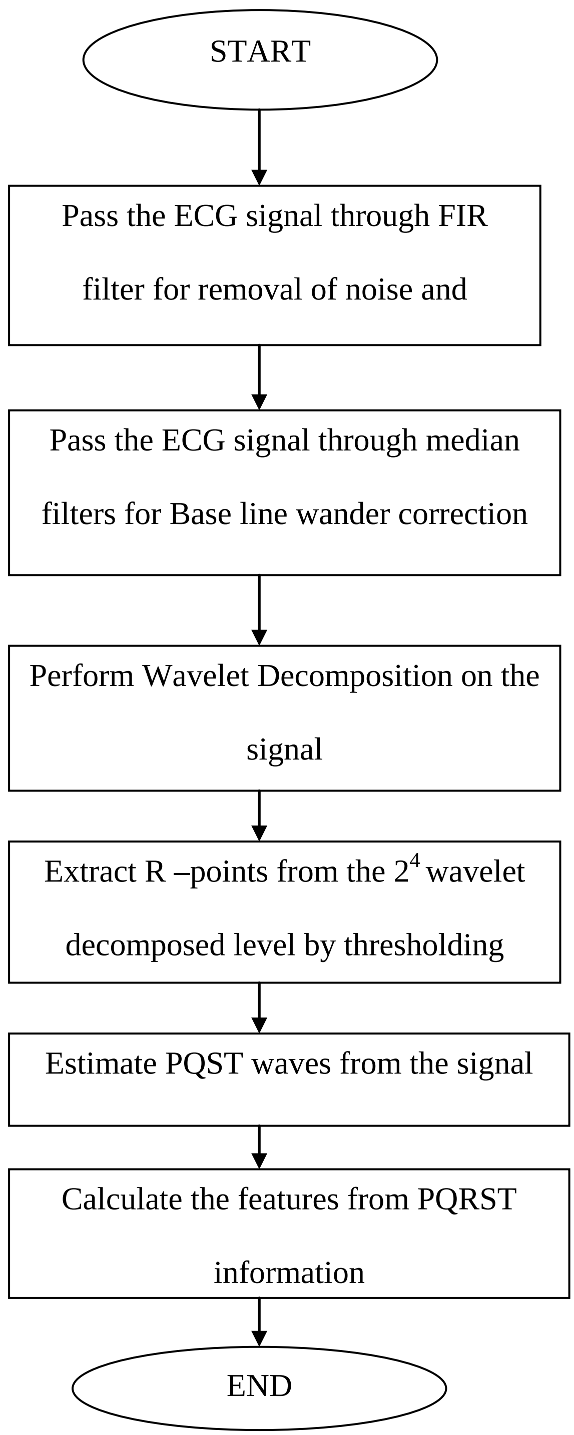

Figure 4. 11: Flow chart for ECG feature extraction

4.8 Classification using Linear Discrimination Analysis

Linear Discriminant analysis is done here also to classify the ECG signal to one of the two groups either seizure or non seizure. LDA was originally used for dimensionality reduction and works by projecting high-dimensional data onto a low dimensional space where the data achieves maximum class separability. In this thesis we are using LDA for classification of ECG signals also. The resulting features in LDA are linear combinations of the original features, where the coefficients are obtained using a projection matrix **W**. The optimal projection or transformation is obtained by minimizing *within-class-distance* and maximizing *between-class-distance* simultaneously, thus achieving maximum class discrimination. The optimal transformation is readily computed by solving a generalized eigenvalue problem.

More formally, for the available samples from the database, we define two measures: (i) *within-class* scatter matrix, given by:

$$S_w = \sum_{j=1}^{M} \sum_{i=1}^{N_i} (\boldsymbol{x}_i^j - \boldsymbol{\mu}_j)(\boldsymbol{x}_i^j - \boldsymbol{\mu}_j)^T \tag{4.2}$$

where $\boldsymbol{x}_i^j$ (dimension $\boldsymbol{nx1}$) is the i^{th} sample vector of class j, $\boldsymbol{\mu}_j$ is the mean of class j, M is the number of classes, and N_i is the number of samples in class j.

The second measure (ii) is called between-class scatter matrix and is defined as:

$$S_b = \sum_{j=1}^{M} (\boldsymbol{\mu}_j - \boldsymbol{\mu})(\boldsymbol{\mu}_j - \boldsymbol{\mu})^T \tag{4.3}$$

where $\boldsymbol{\mu}$ is mean vector of all classes.

The goal is to find a transformation ***W*** that maximizes the between-class measure while minimizing the within-class measure. One way to do this is to maximize the ratio det($\mathbf{S_b}$)/det($\mathbf{S_w}$). The advantage of using this ratio is that if S_w is a non-singular matrix then this ratio is maximized when the column vectors of the projection matrix, ***W***, are the eigenvectors of $\boldsymbol{S_w^{-1}.S_b}$ [**56**]. It should be noted that: (i) there are at most *M-1* nonzero generalized eigenvectors, and so an upper bound on reduced dimension is M-1, and (ii) we require at least *n* (size of original feature vectors) + M samples to guarantee that S_w does not become singular.

In the work discussed here, we use LDA to transform the ECG feature vector of dimension 6 into a reduced feature (of projections) with a varying dimension between 1 and 6. We are using LDA here to classify the features obtained from the above algorithm into two different groups known as seizure and non seizure. The LDA algorithm at first assigns a group to a set of features belonging to the same class. When the algorithm is trained with the set of features available for training it classifies the test vector features to one of the group using Euclidean distance as a measure to know to which group the given signal is belongs to. The LDA is then tested with the evaluate vector for testing the accuracy of the classifier.

4.9 RESULTS AND COMPARISION

We have tested our algorithm with a database of 200 observation of which 100 belong to seizure and 100 belong to non seizure intervals. We have used 45 observation from the seizure and 45 observation from the non seizure to train the LDA. After the LDA is trained with the observation we tested it with 55 observation of seizure and 55 observation of non seizure intervals and found it to correct 93.23% of the time. The variation of accuracy of the algorithm with respect to the features is shown in the figure 4.12 below

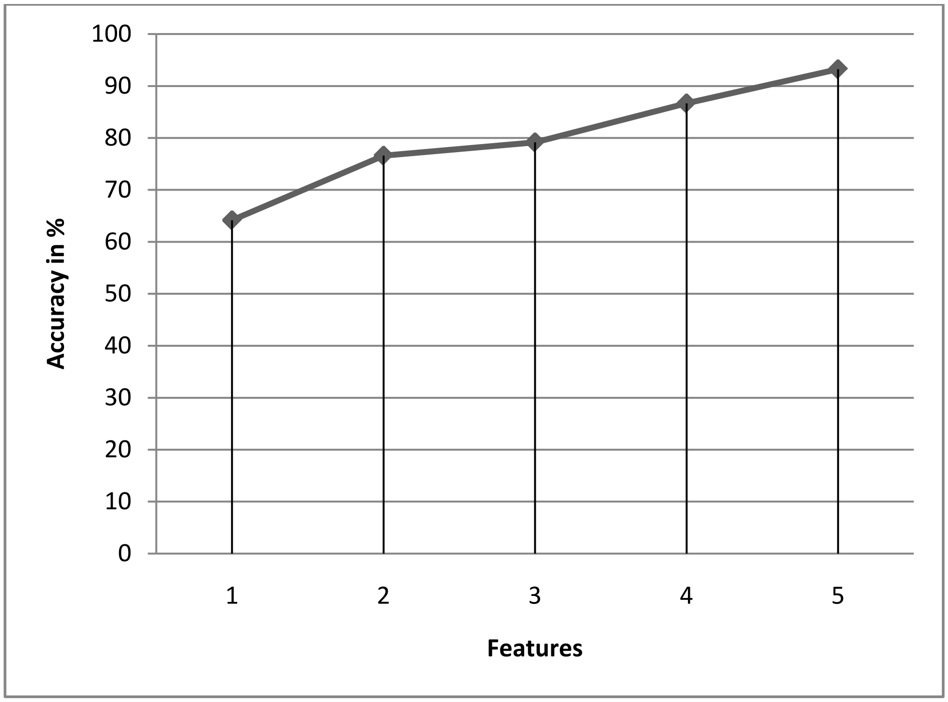

Figure 4. 12: Seizure detection accuracy as a function of the number of features from LDA

The Accuracy, sensitivity and specificity of a classifier are calculated as

$$\text{Accuracy} = \frac{No.\ of\ Correct\ Detection}{Total\ No.of\ Traces\ of\ Healthy\ and\ Seizure\ events}$$

$$\text{Specificity} = \frac{No.of\ True\ Negatives}{No.\ of\ True\ Negative + No.of\ False\ Positives}$$

$$\text{Sensitivity} = \frac{No.of\ True\ Positives}{No.\ of\ True\ Positive + No.of\ False\ Negatives}$$

The specificity of a classifier with 100% means that it identifies all healthy people as healthy whereas a sensitivity of 100% means that it identifies all sick people as sick.

For our classifier we attained a specificity of 96.15% and sensitivity of 98%. The data used in this research is different from the one used by previous researchers. All the research mentioned in the comparison table are done with a different ECG datat set. This is the reason we are presenting a comparison between the sensitivity and specificity measures of the classification algorithms.

Name of the Author of seizure detection using ECG	Accuracy	Sensitiviy	Specificity
D.H.Karim and A.B.Geva	**86%**		
Barry R.Greene	**70.5%**	**62.2%**	**71.8%**
M.B.Malarvili using HRV method		**83.3%**	**100%**
M.B.Malarvili using both time and frequency info.		**85.7%**	**84.6%**
Our technique	**93.23%**	**96.49%**	**90.16%**

Table 4. 1: Performance Comparison

4.10 SECTION SUMMARY

In this section we have presented an algorithm based on ECG signal to effectively classify the given signal into seizure or non seizure event. The ECG features used for classification include R-R mean, R-R variance, P height mean, P height variance, PR duration and QT duration. These features were found to be varying for seizure and non seizure events in the literature. The derived six features are then fed to the LDA for classification which gives an accuracy of 96.37%and specificity and sensitivity of

98.21% and 94.82% respectively. In the next section we are going to discuss about the combination of the seizure detection techniques based on EEG/ECG using Dempster Shafer theory of Evidence.

CHAPTER 5

COMBINATION OF EEG/ECG USING DEMPSTER SHAFER THEORY OF EVIDENCE

5.1 Introduction

The main objective in seizure detection is to achieve highest possible classification accuracy. To attain this objective, many researchers in the past have worked with different combination algorithms. In addition to this different classification algorithms are different in theories, and hence give different amount of accuracy for different applications. "Even though, a specific feature set used with a specific classifier might achieve better results than those obtained using another feature set and/or classification scheme, one cannot conclude that this set and this classification scheme achieve the best classification results"[**79**]. Many combination methods were reported in the past but the important aspect of the combining classifier to be considered is how far the combination method is able to model the uncertainty associated with the performance of each classifier.

5.2 Different approaches for combination of classifiers

The previous researches show that the combination of classifier can be done based on two different ways. The two most important methods for combining the features are:

1. Combination of features (Early integration of classifiers)
2. Combination of classifiers (Late integration of classifiers)

5.2.1 Combination of features (Early integration of classifiers (EI))

In this method the features from ECG and EEG are combined together and fed to the pattern classifier for classification. This method does not need any combination of classifiers as there is only one super feature vector which is the combination of ECG and EEG features. These features are used to train the Linear Discriminant Analysis (LDA) and the classification is based on the Euclidean distance rule to decide which class does the given signal belongs. The figure 5.1 gives the graphical representation of Early Integration of features (EI).

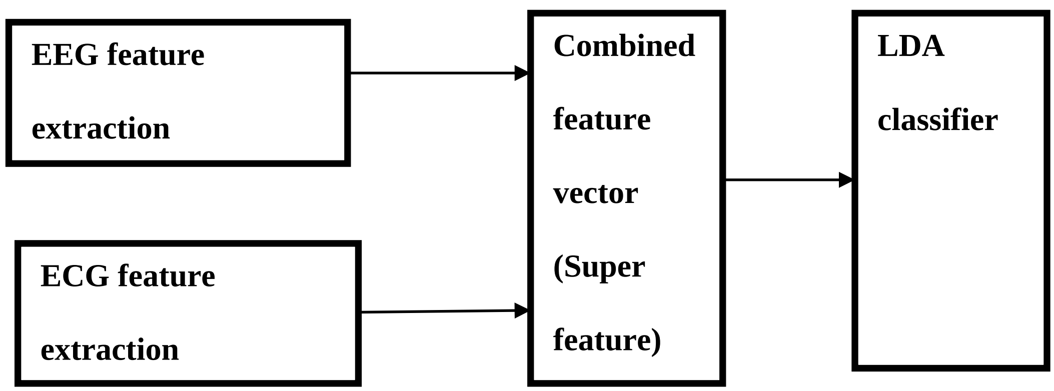

Figure 5. 1: Combination of features (Early Intergration)

5.2.2 Combination of classifiers (Late integration of classifiers (LI))

In this method of classification the individual classifiers are combined instead of features themselves. The features extracted from the ECG and EEG are fed to the LDA for classification and the resulting post probabilities or the decisions are combined using a classifier to get the output result. The figure 5.2 shows the graphical representation of this type of combination

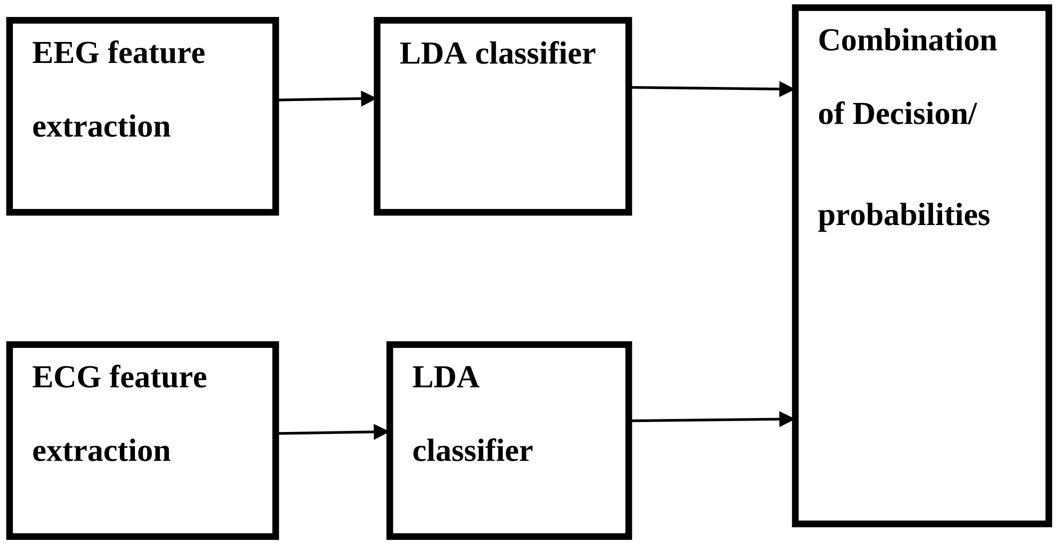

Figure 5. 2: Combination of Classifiers (Late integration)

The combination of classifiers consists of two parts. The first part consists of "How many classifiers are chosen for a specific application and and what kind of classifiers should be used? And for each classifiers what type of features should be used?"[**80**]. Our focus in this chapter is related to the second part of the question which include the problems related to the question "How to combine results of different existing classifiers so that a better result can be obtained ? ".

In the following section we will discuss about the different levels and methods of combination of classifiers.

5.3 Types of Combination of Classifiers

The combination of classifiers can be classified into three types based on the information provided by the output of classifiers.

1. The Abstract level:

   "A classifier only outputs a unique label, or for some extension, outputs a subset"[**80**].

2. The Rank level:

   "A classifier ranks all the labels or a subset of class labels in a queue with the label at the top being the first choice "[**80**].

3. The Measurement level:

   "A classifier attributes to each class a measurement value that reflects the degree of confidence that a specific input belongs to a given class. This degree of belief or confidence could be a single probability value as in a Bayesian classifier or any other scoring measure "[**80**].

5.4 Abstract level Combination

The classifier at abstract level provides the least amount of information and hence is considered as the lowest level of combination. The output of classifier is a single label hence the classifier should be able to provide the abstract output label regardless of the different theories or methodologies the individual classifier may follow. This tye of combination is generally used for all kind of pattern recognition areas. There are many methods of combination discussed at this level. To mention a few popular of them are:

5.4.1 Majority voting

Majority voting is the simplest and most commonly used method for combination of classifiers. "The majority voting system and its variants have achieved very robust and often comparable, if not better, performance than many of the complex system presently available”[**81**]. In simple terms it can be explained as the decision taken by the majority of the classifiers to be taken as the final conclusion result. If *n* classifiers agree to some decision and other set of classifiers less than *n* agree to the other decision then the combination rule assigns the decision in favor of the former one as the majority of classifiers agree with it.

Two basic issues arises during the combination using majority voting which to be summarized are as follows “Should the decision agreed by the majority of experts be accepted without giving due credit to the competence of each expert? Or Should the decision delivered by the most competent expert be accepted, without giving any importance to the majority consensus?”[**81**]. This leads us to the choice between the selection of expert advice or majority consensus based on which there were different majority voting combination schemes presented in the past.

A new method of majority voting which is dependent on the confidences of the individual classifier was presented by L.Lam and C.Y.Suen [**82**] which is called as weighted majority voting. “It is an enhancement to the simple majority system where the classifiers are multiplied by a weight to reflect the individual confidences of the decisions”[**81**]. Further about the weighted majority system is found in [**83**] & [**84**]. There were many variation made in the majority voting later by different researchers. To mention a few are weighted majority voting, class weighted majority voting, restricted majority voting, class wise best decision selection, enhanced majority voting, ranked majority voting , committee methods, regression etc.

5.4.2 Bagging and Boosting

Bagging (Bootstrap aggregating) was proposed in the year 1994 by Leo Breiman [**85**] to improve the combination accuracy of the classifier. “It is a machine learning meta algorithm to improve machine learning and classification and regression models in terms of stability and classification accuracy. It also reduces variance and helps to avoid over fitting. Although it is usually applied to decision tree models, it can be used with any type of model. Bagging is a special case of the model averaging approach”[**86**]. It showed good results in practice but when it comes to weak classifiers, the gains are usually small. An technique for multiple classifier is suitable in these cases known as Boosting.

Boosting deals with the question “whether an almost randomly guessing classifier can be boosted into an arbitrarily accurate learning algorithm. Boosting attaches a weight to each instance in the training set. The weights are updated after each training cycle according to the performance of the classifier on the corresponding training samples. Initially all weights are set equally, but on each round, the weights of incorrectly classified samples are increased so that the classifier is forced to focus on the hard examples in the training set”[**87**].

“There are two major differences between bagging and boosting. First, boosting changes adaptively the distribution of the training set based on the performance of previously created classifiers while bagging changes the distribution of the training stochastically”[**88**]. Second, boosting uses a function of the performance of a classifier as a weight for voting, while bagging uses equal weight voting”[**88**].

5.4.3 Behavior Knowledge Space

Behavior knowledge space is another combination method used at abstract level proposed by Y.S.Huang and C.Y.Suen [**89**]. To avoid independent assumptions, the information is derived from a prior stored knowledge space which records the decision of all classifiers on each learned sample simultaneously[**89**]. The intersection of decisions of each classifier takes one unit of space and for each class the number of incoming samples are accumulated into each unit. The operation of BKS involves two stages “knowledge modeling and decision making. The knowledge modeling stage uses the learning set of samples with both genuine and recognized class labels to construct a BKS. The decision making stage, according to the constructed BKS and the decisions offered from the individual classifiers, enters the focal unit and makes the final decision”[**89**].

5.4.4 Bayesian Formulation

Bayesian combination of classifiers provides the estimates of the posterior probabilities that the given input signal belong to a particular class. A simple Bayesian classification method is given by [**90**].

$$P_{av}\ (X\ \in Wi/X) = \frac{1}{K}\sum_{k=1}^{K} P_k(X\ \in Wi/X)\ ,\ \text{i=1}\ldots M \qquad (5.1)$$

The final classification is done based on the Bayesian criterian,that is the input pattern is assigned to the class to which the posterior probability is maximum.

5.4.5 Dempster Shafer formulation

Dempster Shafer theory was first presented by Arthur P.Dempster and Glenn Shafer in the mid 1970’s , has shown to combine the evidence from different sources. At abstract level it is used to combine the decisions from each classifier and give the degree of belief for the input signal to belong to a particular class. It takes the recognition, substitution, and rejection rates of the classifier to measure the belief of the classifier. When verified experimentally it outperformed majority voting method but the combination at abstract level does not proves to be an optimal combination method as it considers the decisions of the individual classifier instead of their beliefs[**91**].

5.5 Rank level Combination

The output of the classifier at rank level is an ordered sequence of candidate classes, which is called as the n best list. The candidate classes at the first position in the list of classes is considered as the most likely output of the combination classifier and the one at the last of the list is the most unlikely. The candidate classes at the first position is the most likely class, while the class positioned at the end of the list is the most unlikely. Much research is focused on the combination of classifiers at abstract level and measurement level and hence this area is left with very little amount of research in the past[**87**].

5.6 Measurement level Combination

The combination at measurement level has confidence values assigned to each entry of the classifiers. The measurement level combination is the highest level of combination method as the confidence of a classifier gives the useful information which can’t be provided at rank level or abstract level. Most of the research is focused on this combination method as most of the classifiers provide output on this level. To mention few important measurement based combination methods are:

5.6.1 Stacked generalization method

Stacked generalization is a general method of measurement level combination. It works by deducing the outputs of the individual classifier with respect to a provided learning set. “This deduction proceeds by generalizing in a second space whose inputs are the guesses of the original generalizers when taught with part of the learning set and trying to guess the rest of it, and whose output is the correct guess”[**92**]. Different learning algorithm were reported based on this combination method. This was used for regression by Breiman [**93**] and even unsupervised learning by Smyth & Wolpert[**94**][**95**].

5.6.2 Statistical combination method

Different statistical combination methods were discussed by F.Alkoot and J.Kittler [**96**]. The various methods like majority voting, min, max, median etc were compared and the results under normal conditions and disturbed (gaussian noise) were discussed. It was found that the combined classifier gives better results compared to individual classifier especially in the case of median and sum. When Gaussian noise was assumed to be present in the estimation error it was found that single classifier be preferable than product, minimum and maximum[**96**].

5.6.3 Dempster Shafer theory of combination

Dempster Shafer theory of evidence gained much popularity at measurement level. The theory is a generalization of Bayesian formulation. This theory introduced the system of beliefs in the output results which were not found to be discussed in the previous combination techniques and hence it gained attention by the researchers as it gave a meaningful reason for the combined result.“It was adopted in Artificial Intelligence by researchers in order to process probabilities in expert systems, but has soon been adopted for other application areas, such as sensor fusion and classifier combination”[**87**]. More about the DST will be discussed after discussing the problem related to uncertainty and the use of DST to be an appropriate approach when it comes to representing uncertainty.

5.7 Problem of Uncertainty

Recently the researchers are focused on the importance of modeling uncertainty. The two types of uncertainty generally associated with any system are classified as follows

1. Aleatory Uncertainty:

   The type of uncertainty which results due to the fact that the system can behave in random ways (ex: Noise)[**97**].

2. Epistemic Uncertainty:

   The type of uncertainty which results from the lack of knowledge about a system and is a property of the analysts performing the analysis and hence this type of uncertainty is a Subjective uncertainty[**97**].

The first type of uncertainty is generally overcome by using the frequentist approach associated with traditional probability but the problem is with the second type of uncertainty which represents the lack of knowledge related to some event. In the probability theory it is necessary to have the knowledge on all types of events. When this is not available uniform distribution function is often used, which means that all simple events for which a probability distribution is not known in a given sample space are equally likely. An additional axiom of the Bayesian theory is that the sum of the belief and disbelief in an event should add to 1 i.e. P (x) + P $(\bar{x})$ =1.The D-S theory of evidence rejects this axiom outwardly and introduces the concept of beliefs and allows the combination of evidence obtained by multiple sources and the modeling of conflicts between them.

Let us further explain the above statements with an example to clear the concept of uncertainty. Suppose $\boldsymbol{\Phi}$ represents a statement: the place is beautiful. Then according to the classical theory of Bayesian the theorem P($\boldsymbol{\Phi}$) + P($\bar{\Phi}$) =1 , where $\bar{\Phi}$ represents then negation of the proposed statement. Now consider a person x who has not ever visited the place at all and thus he does not have any idea about how the place looks like and also he cannot say that he does not belief in the above statement. Here comes the concept of uncertainty and a limitation to the Bayesian theory. This concept is well explained by the use of Dempster Shafer theory. The Dempster Shafer theory notes down the belief of the person x in the given statement m($\boldsymbol{\Phi}$)=0 and disbelief m($\bar{\Phi}$)=0 indicating that the person x is uncertain of the event.

Thus the major difference between the Bayesian formulation and Dempster Shafer theory in solving is conceptual. The statistical model assumes that there exist Boolean phenomena where as the D-S theory concerns for the belief in that particular event. “The result of the Bayesian formulation leads to the assumption that commitment in belief of a certain hypothesis leads to the commitment of the remaining belief to its negation. Thus if we belief in the existence of certain hypothesis this would imply, under the Bayesian formulation a large belief to it non existence, which is what we call over commitment. In D-S theory one considers the evidence in favor of hypothesis. There is no causal relationship between a hypothesis and its negation. Rather, lack of belief in any particular hypothesis implies belief in the set of all hypotheses, which is referred to as the state of uncertainty. If the uncertainty is denoted by θ then for the above example m(θ)=1, which is calculated by the following formula: $m(\boldsymbol{\Phi}) + m(\bar{\Phi}) + m(\theta) = 1$”[**91**].

This is the reason for selecting the D-S theory as combination rule in our thesis. In the following section we are going to discuss about the basic concepts of D-S theory.

5.8 Dempster Shafer Theory of Evidence

The Dempster Shafer theory was introduced by Glenn Shafer and A.P.Dempster as a generalization of Bayesian theory. It is famously known as the theory of belief functions. It is a very powerful technique when it comes to modeling uncertainty. “An important aspect of this theory is the combination of evidences obtained from multiple sources and modeling the conflict between them”[**98**]. It is usually based on two main ideas: the first being the idea of obtaining belief function of one’s degree of belief and the second being the reasoning mechanism involved on the combination rule.

We now present 3 basic concepts related to D-S theory. They are

1. Basic belief assignment
2. Belief function
3. Plausibility

5.8.1 Basic belief assignment (BBA)

A basic belief assignment is (bba) *b(.)* is the basic of evidence theory. It assigns a value between 0 and 1 to all the variables in the subset A where the bba of the null set is

0 and the summation of bba's of all the subsets and should be equal to 1. This is given as follows:

$$b(\varphi) = 0, and \ \sum_{A \subseteq \theta} b(A) = 1 \qquad (5.2)$$

Where φ is a null set. The bba b(.) for a given set U represents the amount of belief that a particular element of X (universal set) belongs to the set U (represented by m(A)) but to no particular subset of A. The value of b(A) pertains only to set U and makes no additional claims about any subsets of A. Any further evidence on the subsets of A would be represented by another bba b(B), where B is a subset of A[**98**].

5.8.2 Belief function

The belief function is used to assign a value [0,1] to every nonempty subset B. For every probability assignment two bounds of intervals can be defined. The lower bound in the case of D-S theory is represented by belief function. It is defined as the sum of all the basic belief assignments bba's of the proper subsets of (B) of the set of interest (A) (B⊆A). It is called as degree of belief in B and is defined by

$$\text{Bel (A)} = \sum_{B \subseteq A} b(B) \qquad (5.3)$$

where B is a subset of A. The belief function can be considered as a generalization to probability distribution function whereas the basic belief assignment can be considered as a generalization to probability density function[**91**].

5.8.3 Plausibility

The upper limit of the probability assignment is called as plausibility. It is the sum of all the probability assignments of the sets (B) that intersect the set of interest (A) ($B \cap A \neq \boldsymbol{\Phi}$).

$$Pl(A) = \sum_{B/B\cap A\neq\varphi} b(B) \tag{5.4}$$

The belief and plausibility measures represent the lower and upper bound of probability for a given hypothesis. These two are non additive as the sum of all belief functions or the sum of all plausibility functions need not be necessarily equal to 1.

5.8.4 Combination rule

The combination rule in D-S theory depend on the basic belief assignments b(.). Let b_1(.) and b_2(.) be two basic belief assignments for the belief function bel_1(,) and bel_2(.) respectively and these two belief functions are the focal element of the set B_j and C_k respectively. Then the combine belief commited to $A\subseteq\theta$ is given by

$$b_{12}(A) = \frac{\sum_{B\cap C=A} b_1(B)b_2(C)}{1-K} \text{ when } A\neq\emptyset \tag{5.5}$$

Where $K=1-\sum_{B\cap C=\emptyset} b_1(B)b_2(C)$

The denominator K here represents s the basic probability mass and is associated with conflict. The whole term 1-K represents the normalizing factor which has the effect of completely ignoring the effect of conflict and attributing any probability mass associated with conflict to the null set[**99**]. The above theory of Dempster Shafer can be well explained by understanding an example below.

5.9 Example

Consider an example of a car parked in a parking lot. Say now Jack comes to the office and says that the car is not there. But we know that the Jack is absent minded and hence he is correct only 80% of the time. Suppose now another person Jill comes to the office and says the same thing but we know that Jill is correct only 70% of the time. From this available information we will calculate the beliefs of each.

As we know that the Jack is correct only 80% of the time and thus the evidence for the car missing in the lot is 80% and for the rest 20% we don't have any information one way or the other. Hence we can say that the probability of the car missing in the lot is 0.8 and might be up to 1.0. This is what we call a probability interval [0.8 1.0]. Instead of having one definite value for calculating the probability we have captured the information by a probability interval. The lower bound in the interval is called as belief and the upper bound is called as plausibility. The two can be related as given in the equation below

$$\text{Bel}\ (p)=1\text{-Pl}(\bar{p}) \tag{5.6}$$

Bel(p) shows how certain we are about missing the car, where as the second term indicates how much high can be the probability of missing the car given how certain we are about being the car in the correct place. As the evidence of car being in the correct place is zero and hence the plausibility of the event of the car being missed will be equal to 1.0.

Similarly the probability interval for the belief of Jill can will be [0.7 1.0]. Now if we want to combine the evidences the combined probability of that both Jack and Jill are unreliable will be 0.3*0.2=0.06. It means that the information about the car being missing is 94% correct. So, now the new belief is 0.94 and the interval is [0.94 1.0]. In this case we considered that both of them were consistent in the evidence of car being missed. Now if we consider a case where Jack says that the car is missing and Jill says that it is there. Thus the new probability intervals for Jack and Jill would be [0.8, 1.0] & [0, 0.3] respectively. We will have four different cases now

1. Both Jack and Jill are reliable, impossible as both cannot be correct at the same time.
2. Jack is reliable and Jill is not, with probability 0.8* 0.3=0.24. The car will be missing in this case.
3. Jill is reliable and Jack is not, with probability 0.2*0.7=0.14. The car will be present in this case.
4. Both of them are unreliable, with probability 0.2*0.3=0.06. The information will be uncertain in this case.

In order to convert this probability information into beliefs we have to normalize.

We know by Dempster Shafer rule the sum of three probabilities should be equal to one i.e $m(\boldsymbol{\Phi}) + m(\bar{\Phi}) + m(\theta) = 1$. But, if we sum up the above three probabilities it will be equal to 0.24+0.14+0.06= 0.44 and this is not equal to 1. So to normalize the above probabilities we have to divide the probabilities by 0.44, thus the probability of a missing car will be 0.24/0.44= 0.545 and the car to be present will be 0.14/0.44=0.318. The possibility interval for the car being missed will be then [0.545, 1-0.318] which equals [0.545 0.682]. The lower bound is the belief function and the upper bound is the plausibility.

Thus in this way we will be calculating of the beliefs and plausibility. The combination of the results is done according to the Dempster Shafer equation given by equation 5.5 . This combination technique is used for combining the results obtained from ECG and EEG for classifying the results to belong to one of the two classes, viz seizure and non seizure.

5.10 Dempster Shafer combination Algorithm

The Combination of Results from both the classifiers is done using the Dempster Shafer Rule. For this the information available to us from the ECG/EEG algorithms should be in the form of probability information. The Step by Step algorithm for combining the results using Dempster Shafer theory of evidence is discussed below:

Step 1: Calculating the Normalized distance

The first and foremost thing to be done before extracting the beliefs is to extract the probability information from the ECG/EEG algorithms. For this the Euclidean distance between the feature vector under test and the mean of the seizure class feature vectors and non seizure class vectors is calculated as shown in equation.

$$v = \frac{x-\mu}{\sigma} \qquad (5.7)$$

Where x = Test feature vector

μ = Mean of the Class feature vectors

σ=variance of the Class feature vectors

Step 2: Extracting the Probability information

The value obtained in equation is substituted in the normal distribution to get the probability value for seizure and the probability value of non seizure of an event.

Step 3: Assignment of Basic Belief

From the probability information the Basic Belief is calculated. The probability of a seizure event is assumed as the Belief in seizure event and the probability of normal case is considered as Belief in non seizure. The conflict between the two probability values is considered as the Uncertainty of information.

Step 4: Belief and Plausibility

From this Basic Belief the Belief and Plausibility of the event is calculated. This is calculated using the equation 5.8. The Belief represents the minimum probability of happening of an event and plausibility represents the maximum amount of probability of happening of the event.

$$\text{Bel}\ (p)=1\text{-Pl}(\bar{p}) \qquad (5.8)$$

Step 4: Combining the Beliefs using Dempster Shafer Rule

The resulting belief functions are then combined using the Dempster Shafer Theory as follows

$$b_{12}(A) = \frac{\sum_{B\cap C=A} b_1(B)b_2(C)}{1-K} \text{ when } A\neq \emptyset \qquad (5.5)$$

$$\text{Where K}=1-\sum_{B\cap C=\emptyset} b_1(B)b_2(C)$$

Where 1-K represents the normalizing factor

Step 5: Thresholding

The resultant belief is then threshold by a value of 0.5. This method of thresholding is done to classify the results to belong to any one of the class viz seizure and non seizure events.

Flow Chart for Combination Algorithm:

The Flow Chart for the above algorithm is shown in the figure 5.3 below

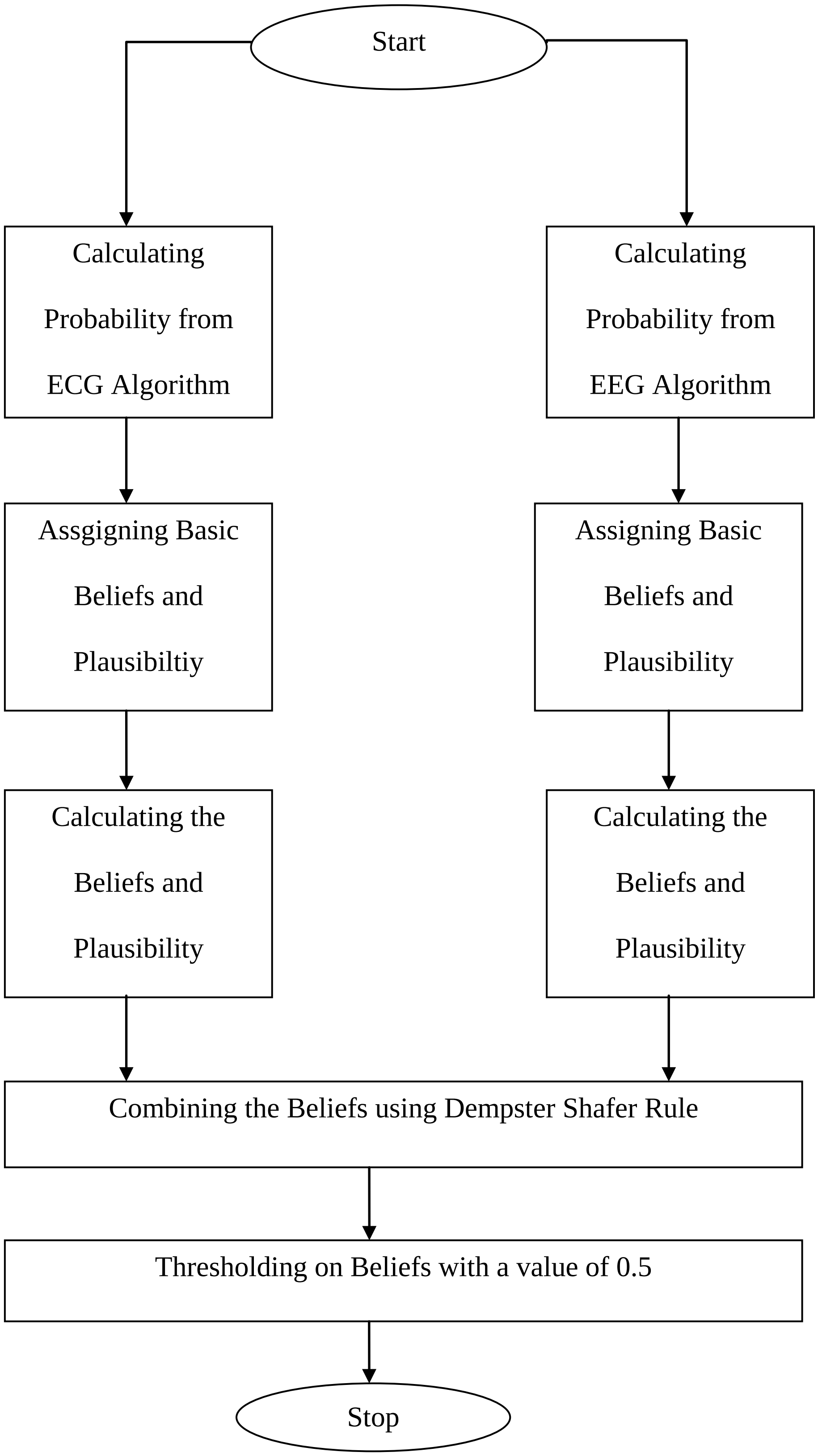

Figure 5. 3: Flow Chart for Combining results of ECG/EEG using Dempster Shafer theory of Evidence

5.11 Combined classification result

In this section we are going to discuss the results of D-S theory under two different cases.

Case 1:

Here we take the healthy traces and seizure traces and train the LDA to recognize healthy traces as belonging to group1 and seizure traces to group2 for both EEG and ECG algorithm. Now the individual classifiers are combined using Dempster Shafer theory using the above algorithm.

We have used 90 traces of EEG and ECG for training the LDA and 110 traces for testing. When the results of each classifier were combined using D-S theory of combination we achieved an accuracy of 95.57%. The results are compared in table 2.

	Accuracy	Sensitivity	Specificity
ECG	93.23%	96.49%	90.16%
EEG	90.00%	92.50%	89.20%
D-S combination of EEG and ECG	96.90%	94.71%	94.90%

Table 5. 1: Combination of EEG, ECG & D-S combined algorithm (CASE 1)

Case 2:

Now we add 5 traces of healthy and 5 traces of seizure to the individual ECG/EEG algorithm and mention it as to belong to group 3. The classification algorithm should be able to classify the results to belong to either class 1 or class 2. This causes reduction in the accuracy of the individual classifiers. The accuracy of the seizure detection algorithm for EEG and ECG now drops to 84.16% and 75.83% respectively. Now if we use the Dempster shafer theory of evidence for combining the classifiers it gives an average accuracy of 90.74%.

	Accuracy	Sensitivity	Specificity
ECG	75.83%	78.94%	82.19%
EEG	84.16%	86.95%	84.50%
D-S combination of EEG and ECG	90.74%	93.64%	92.89%

Table 5. 2: Comparison of ECG,EEG & D-S Combination algorithms (CASE 2)

Receiver Operating Characteristics (ROC) :

ROC curve is mainly used in signal processing theory to provide optimal models and to discard suboptimal ones. It is used as a statistical tool for measuring the robustness of the classifier. It is a plot of the Sensitivity Vs 1-Specificity or true positive rates vs false positive rates by varying the threshold of the classifier. The ROC for case 1 and case 2 are shown in the figure 5.4 and 5.5 below respectively. It was found that ROC for case1 has an area of 95.35% under the curve and the ROC for case 2 has an area of 92.85% to give under the curve.

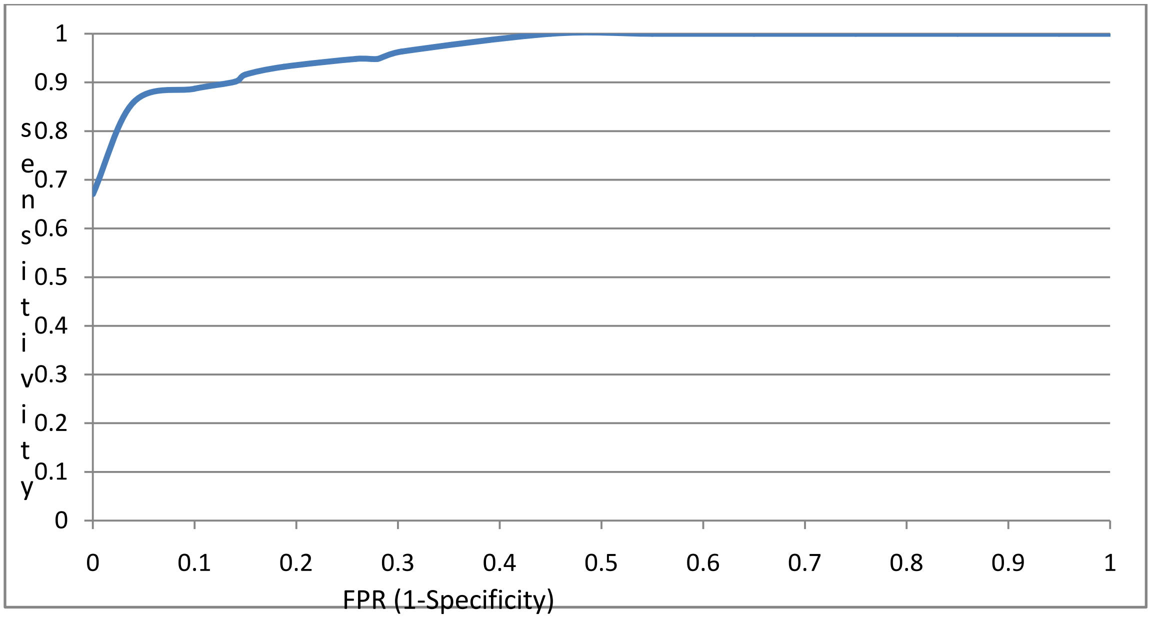

Figure 5. 4: Receiver Operating Characteristics (ROC) for Case 1

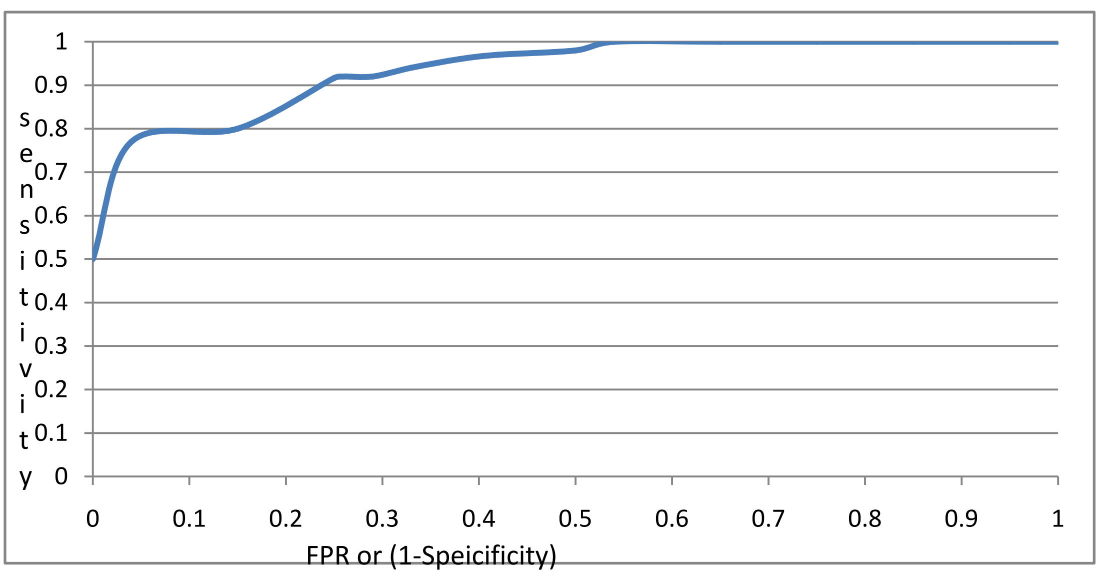

Figure 5. 5: : Receiver Operating Characteristics (ROC) for Case 2

5.12 Degree of Association

The data used for EEG and ECG in this research belong to different databases. So in order to show the degree of association between the two different databases we performed a small test.

We have a database of 90 ECG/EEG traces for testing and 110 ECG/EEG for training. We assume x persons ECG to belong to y^{th} person's EEG. To show the degree of association we shift 10 samples of EEG database each time and associate it with the ECG database. At each shift we measure the detection accuracy of the algorithm. The effect of this shift on the combination accuracy for case 1 and case 2 are shown in the tables 5.3 and 5.4 below.

Shift	Accuracy	Sensitivity	Specificity
1^{st}	94.16%	96.70%	92.30%
2^{nd}	94.62%	95.20%	94.48%
3^{rd}	96.68%	98.30%	95.20%
4^{th}	95.83%	96.70%	95.23%
5^{th}	97.24%	98.36%	96.77%
6^{th}	95.00%	96.70%	93.70%
7^{th}	94.16%	95.20%	93.75%
8^{th}	97.45%	98.30%	96.74%
$S9^{th}$	93.33%	95.23%	92.30%
10^{th}	97.24%	98.36%	96.77%

Table 5. 3: Degree of Association for Case 1

Shift	Accuracy	Sensitivity	Specificity
1st	95.00%	95.23%	95.23%
2nd	90.83%	90.90%	92.30%
3rd	92.50%	93.70%	90.90%
4th	91.66	92.30%	92.30%
5th	93.33	93.7%	93.70%
6th	95.83	95.23%	96.70%
7th	92.50%	95.23%	90.90%
8th	90.83%	90.90%	92.30%
9th	91.66%	93.75%	90.90%
10th	93.33	93.7%	93.70%

Table 5. 4: Degree of Association for Case 2

It can be seen from the tables that for case 1 the average accuracy was found to be 95.57% and the standard deviation to be 3.91%. From the case 2 it can be seen from the table 5.4 that the average accuracy is 90.747% and the standard deviation to be 4.17%.

5.13 Summary

In this chapter we have discussed about various combination techniques for combining the results obtained for EEG and ECG algorithms. It was found in the research that Dempster Shafer theory of evidence is best suited when it comes to modeling uncertainty while combining the belief of different classifiers. The individual classifiers are then combined using Dempster Shafer theory of Evidence. The results obtained for the D-S theory for different cases are observed and found that the combination of EEG & ECG algorithms using D-S theory gives good results.

CHAPTER 6

FUTURE WORK AND CONCLUSIONS

In this thesis we have designed a robust seizure detection technique which can detect seizure even in the presence of uncertain information from any of the inputs.

We have designed a time frequency based seizure detection technique which uses the EEG signal and extracts the left singular values from the time frequency matrix of the EEG signal to train the LDA. The different types of time frequency representation of EEG signal are discussed and Wigner ville distribution is selected to represent the EEG signal in time frequency domain as it is giving sharp features related to seizure trace of EEG signal. The result of the TF-LDA algorithm gives an average accuracy which outperforms the previously mentioned seizure detection algorithms.

We have designed a seizure detection algorithm based on ECG which considered the features from the ECG wave for seizure detection which were not utilized in the past for detection of seizures. The ECG features used for classification include R-R mean, R-R variance, P height mean, PR duration and QT duration. The derived five features are then fed to the LDA for classification. These features were found to give good classification accuracy with good specificity and sensitivity rates.

Finally we combined both algorithms using Dempster-Shafer theory of evidence. It was found in the research that Dempster-Shafer theory of evidence is best suited when it comes to modeling uncertainty while combining the belief of different classifiers. The individual classifiers are then combined using Dempster-Shafer theory of Evidence. We have tested the combination under two different cases.

1. In the case 1 we take the healthy traces and seizure traces and train the LDA to recognize healthy traces as belonging to group1 and seizure traces to group2 for both EEG and ECG algorithm. Now the individual classifiers are combined using Dempster Shafer theory using the above algorithm. The results obtained gave accuracy better than the individual classifiers.
2. In the case 2 we added 5 traces of healthy and 5 traces of seizure to the individual ECG/EEG algorithm and mention it as to belong to group 3. The classification algorithm should be able to classify the results to belong to either class 1 or class 2. This resulted in reduction in accuracy of the individual classifiers. Now if we use the Dempster-shafer theory of evidence for combining the classifiers it gives an average accuracy comparable to the case 1 which shows that the Dempster Shafer theory of combination is a robust combination technique which can give good results even in the presence of uncertainty of information.

6.1 Future Work

The following are the recommendations for future work in this field

- In addition to the above method we can increase the accuracy by using the combination of more than 2 methods for detecting seizures based on ECG or EEG.
- The different combination schemes can be done at abstract or measurement level.

- Robustness can be improved by considering the effects of seizure on Respiration rate and Body movements and using the combination of all different methods of recognizing seizure.
- Electrocorticography (ECoG) is a method of recording the brain activity by placing the electrodes on the surface of brain. Future work in this field for automatic seizure detection is yet to be covered.

References

[1] P.L.Paige and P.R.Carney, *Neurological Disorders, Handbook of Neonatal Intensive care*. USA: St. Louis , 2002.

[2] C.T.Lombroso, *Neonatal EEG polygraphy in normal and abnormal newborns in Electroencephalography Basic principles Clinical applications and Related fields*. Baltimore, Md , USA, 1993.

[3] Volpe JJ, *Neurology of the newborn*. Philadelphia :Saunders, 2001.

[4] Aso K, Beggarly ME, Hamid MY, Steppe DA, Painter MJ, Scher MS, "Electrographic seizures in preterm and full term neonates: clinical correlates, associated brain lesions, and risk for neurologic sequale," 1993.

[5] MD Bola Adamolekun, "Seizure Disorders," *The Merck Manuals online medical library*, March,2008.

[6] Vestergaard M, Mortensen PB, Sidenius P, Agerbo E Christensen J, "Epilepsy and risk of suicide: a population-based case-control study," Aug 2007.

[7] J.S.Hahn, G.P.Heldt and R.W.Coen A.Liu, "Detection of neonatal seizures through computerized EEG analysis," vol. vol.1, 1992.

[8] National Institure of Neurological Disorders and Stroke, "http://www.ninds.nih.gov/disorders/epilepsy/detail_epilepsy.htm#175223109," USA,.

[9] Linda J. Vorvick, "Seizures," *National Institutes of Health"http://www.nlm.nih.gov/medlineplus/ency/article/003200.htm"*, March,2010.

[10] Retrieved from Epilepsy.com, "What is Epilepsy?," *http://www.epilepsy.com/pdfs/what_is_a_seizure.pdf*.

[11] "http://www.ehealthmd.com/library/epilepsy/EPI_whatis.html,".

[12] Hsun-Hsien Chang and Jose M.F.Moura, *Biomedical Engineering and Design Handbook*.: Tata Mc Graw Hill, 2010.

[13] Niedermeyer E. and da Silva F.L., *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*.: Lippincot Williams & Wilkins, 2004.

[14] J.S.Hahn, G.P.Heldt and R.W.Coen A.Liu, "Detection of neonatal seizures through computerized EEG analysis," vol. vol.1, 1992.

[15] D.Flangan, B.Rosenblatt, A.Bye and E.M.Mizrah J.Gotman, "Evaluation of an automatic seizure detection method for the newborn EEG," 1997.

[16] Frei MG, Wilkinson SB Osorio I, "Real time automated detection and quantitative analysis of seizures and short term prediction of clinical onset," June 1998.

[17] Paul Colditz Patrick Celka, "A Comuter-Aided Detection of EEG Seizures in Infants: A Singular-Spectrum Approach and Performance Comparision," vol. 49, no. 5, May 2002.

[18] T.He, L.A.Smith, and L.Tarassenko P.E.McSharry, "Linear and non linear methods for automatic seizure detection in scalp electro encephalogram recordings," 2002.

[19] Guy Dumont, Donald Gross, Craig R. Ries, Ernie Puil, and Bern A.MacLeod Reza Tafreshi, "Seizure detection by a novel wavelet packet method," Aug.2006.

[20] Mim Lim Choo, U.Rajendra Acharya, P.K.Sadasivan N.Kannathal, "Entropies for detection of epilepsy in EEG," vol. 80, no. 3, April 2005.

[21] Abdulhamit Subasi, "Automatic detection of epileptic seizure using dynamic fuzzy neural networks," 2006.

[22] M.Mesbah and B.Boashash H.Hassanpour, "Time frequency based new born EEG seizure detection using low and high frequency signatures," 2004.

[23] H. Carson, and M. Mesbah B. Boashash, "Detection of seizures in newborns using time-frequency analysis of EEG signa," 2000.

[24] Samanwoy Ghosh-Dastidar, and Nahid Dadmehr Hojjat Adeli, "A Wavelet-Chaos Methodology for Analysis of EEGs and EEG subbands to detect seizure and epilepsy," vol. 54.

[25] Reinhard Grebe, Fabrice Wallois Ardalan Aarabi, "A multistage knowledge based system for EEG seizure detection in newborn infants," vol. 118, no. 12, Dec. 2007.

[26] Min Soo Kim and Hee Don Seo Berdakh Abibullaev, "Seizure detection in temporal lobe epileptic EEGs using the best basis wavelet functions," feb. 2009.

[27] Rakesh Kumar Sinha, Rajesh Hatwal, Barda Nand Das Anup Kumar Keshri, "Epileptic spike recognition in electro encephalogram using deterministic finite automata," vol. 33, no. 3, June 2009.

[28] Javidan M, Dumont GA, Tafreshi R Zandi AS, "Automated real time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform," July 2010.

[29] R.J. Vermeulen, R.L.Strijers, W.P.Fetter and C.J.Stam J.Altenburg, "Seizure detection in the neonatal EEG with synchronization likelihood," 2003.

[30] H.Hassanpour,M.Mesbah L.Rankine, "Newborn EEG simulation from Non-linear analysis," 2005.

[31] H.Otsubo, S.Parvez, A.Lodha, E.Ying, B.Parvez, R.Ishii, Y.Mizuno-Matsumoto, R.A.Zoroofi and O.C.Snead M.Kitayama, "Wavelet analysis for neonatal electroencephalographic seizures," 2003.

[32] Newborns EEG seizure detection using a time frequency approach, "Pegah Zarjam and Ghasem Azemi".

[33] A.B.Geva D.H.Kerem, "Forecasting Epilepsy from the Heart Rate Signal," 2005.

[34] Philip de Chazal, Geraldine B.Boylan, Sean Connolly, and Richard B.Reilly Barry R.Greene, "Electrocardiogram Based Neonatal Seizure Detection," vol. 54, April 2007.

[35] Mostefa Mesbah, and Boualem Boashash M.B.Malarvili, "Time Frequency Analysis of Heart Rate Variability for Neonatal Seizure Detection," 2007.

[36] M.B.Malarvili and Mostefa Mesbah, "Newborn Seizure Detection Based on Heart Rate Variability," vol. 56, Nov.2009.

[37] http://en.wikipedia.org/wiki/Electrocorticography,.

[38] Mark G. Frei, Jon Giftakis, Tom Peters, Jeff Ingram, Mary Turnbull, Michele Herzog, Mark T.Rise, Scott Schaffner, Richard A.Wennberg, Thaddeus S.Walczak, Michael W. Risinger, and Cosimo Ajmone-Marsan Ivan Osorio, "Performance reassessment of a real time seizure detection algorithm on long ECoG series," 2002.

[39] G.Tao, J.Frost Jr., M.Wise, R.Hrachovy, E.Mizrahi N.Karyiannis, "Automated detection of videotaped neonatal seizures based on motion segmentation methods," 2006.

[40] Geraldine B.Boylan, Richard B.Reilly, Philip de Chazal, Sean Connolly Barry R.Greene, "Combination of EEG and ECG for improved automatic neonatal seizure detection," 2007.

[41] David Lowe and Anne-Marie Arlaud-Lamborelle T.Bermudez, "Multimodal model fusion of EEG/ECG for epileptic seizure detection," 2007.

[42] David Lowe and Anne-Marie Arlaud-Lamborelle T.Bermudez, "Schemes for

fusion of ECG and EEG towards temporal lobe epilepsy diagnostics," IEEE, 2007.

[43] Srinivasan R Nunez PL, "Electric fields of the brain: The neurophysics of EEG.," 1981.

[44] S., Thorne, B. M. Klein, *Biological psychology*. New York, 3 October 2006.

[45] B. Abou-Khalil and K.E. Musilus, "Atlas of EEG & Seizure Semiology," 2006.

[46] http://en.wikipedia.org/wiki/File:EEG_cap.jpg,.

[47] Klaus Lehnertz, Florian Mormann, Chritoph Rieke, Peter David, and Christian E.Elger Ralph G.Andrzejak, "Indications of nonlinear deterministic and finite-dimensional stuctures in time series of brain electrical activity: Dependence on recording region and brain state," vol. 64, no. 061907, Nov. 2001.

[48] Lijie Yu, Haoyuan Gao Ye Yuan Yue Li, "Analysis of non linearity in normal and epileptic EEG signals,".

[49] Jean Jacques Bellanger, Fabrice Bartolomei, Fabrice Wendling, and Lotfi Senhadji Karim Ansari-Asi, "Time frequency characterization of interdependencies in nonstationary signals: application to epileptic EEG," July 2005.

[50] A.Papandreu-Suppappola, *Applications in time frequency signal processing*.: CRC press, Arizona, 2003.

[51] L.Cohen, "Generalized phase space distribution functions," vol. 7, 1966.

[52] H. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," vol. 37, no. 6, June 1989.

[53] L. E. Atlas, and R. J. Marks Y. Zhao, "The use of cone-shape kernels for generalized time-frequency representations of nonstationary signals," vol. 38, no. 7, July 1990.

[54] B. Boashash, "Time-Frequency Signal Analysis and Processing: A Comprehensive Reference," 2003.

[55] Aliex M.Martinez & Avinash C.Kak, "PCA Vs LDA".

[56] J. Belhumeur, P. Hespanha, D. Kriegman N. Peter, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," July, 1997.

[57] S.Wiebe, W.T.Blume, R.S.McLachlan, G.B.Young, S.Matijevic C.Deacon, "Seizure identification by clinical description in in temporal lobe epilepsy," july 2003.

[58] Cormilas J, Hirsch LJ Opherk C, "Heart rate and EKG changes in 102 seizures: analysis of influencing factors," vol. 52, no. 2, December 2002.

[59] Reginald T.Ho, and Michael R.Sperling Maromi Nei, "EKG Abnormalities during partial sezures in refractory epilepsy," 2000.

[60] Cardiovascular Consultants, "Heart Information Center," 2006.

[61] Hall JE Guyton AC, *Textbook of Medical Physiology*.: WB Saunders Co, 1996.

[62] Auseon JC, Waksman D Brose JA, "The Guide to EKG Interpretation White Coat Pocket Guide Series," 2000.

[63] Edward Labrador Kasturi Joshi, "Early Myocardial Infraction Detection," 2009.

[64] http://en.wikipedia.org/wiki/File:SinusRhythmLabels.svg,.

[65] Hall WD, Hurst JW Walker HK, *Clinical methods: the history, physical, and laboratory examinations (in English)*., 1990.

[66] Richard E.Klabunde, *Cardiovascular Physiology Concepts*.: Lippincott Williams & Wilkins, 2005.

[67] P.E.M.Smith and Lynne Owen L.D.Blumhardt, "Electrocardiographic Accomplishment of Temporal Lobe Epileptic Seizures," vol. 327, no. 8489, 1986.

[68] Christiana Schernthaner, Stefanie Lurger, Klaus Potzelberger, Christoph Baumgartner Fritz Leutmezer, "Electrocardiographic Changes at the Onset of Epileptic Seizures," vol. 44, no. 3, March 2003.

[69] Stephen Oppenheimer, "Cardiac dysfunction during seizures and the sudden epileptic death syndrome," 1990.

[70] "Cardiac Arrest during Seizure," 2000.

[71] Carson R Reider, Amparo Kay Miles E.Drake, "Electrocardiography in epilepsy

patients without cardiac symptoms," vol. 2, March 1993.

[72] Marchi Stoshak and Timothy J.Rittenberyy Linda L.Herman, "Long QT Syndrome presenting as a seizure".

[73] Sudden Unexpected Death in Epilepsy (SUDEP), , August 2010.

[74] KB Krishnamurthy, JM Hausdorff , JE Mietus, JR Ives, AS Blum, DL Schomer, AL Goldberger IC Al-Aweel, "Post Ictal Heart Rate Oscillations in Partial Epilepsy," 1999.

[75] MATLAB 7.0,.

[76] Robi Polikar, "http://users.rowan.edu/~polikar/WAVELETS/WTpart3.html".

[77] D.C.Reddy N.Sivannarayana, "Biorthogonal wavelet transforms for ECG parameters estimation," Dec. 1998.

[78] Tompkins Willis.J,.: Prentice Hall, 2000.

[79] G.Giacinto, *Design of multiple classifier systems*.: University of Salerno, 1998.

[80] Adm Krzyzak and Ching Y.Suen Lei Xu, "Methods of combining multiple classifiers and their application to handwriting recognition," vol. 22, no. 3, June 1992.

[81] H.Alam and M.C.Fairhurst A.F.R.Rahman, "Multiple classifier combination for character recognitio: Revisiting the majority voting system and its variations".

[82] L.Lam and C.Y.Suen, "A theorotical analysis of the application of majority voting to pattern recognition," , 1994.

[83] Y.S.Huang, and C.Y.Suen L.Lam, *Combination of multiple classifier decisions for optical character recognition in Handbook of character recognition and document image analysis,pages 79-101*.: World Scietific Publishing Company, 1997.

[84] A.Hojjatoleslami and T. Windeatt J.Kittler, "Weighting factors in multiple expert fusion," , 1997.

[85] Leo Breiman, "Bagging Predictors," Berkely, CA, 1996.

[86] http://en.wikipedia.org/wiki/Bootstrap_aggregating,.

[87] Stefan Jaeger, and Venu Govindaraju Sergey Tulyakov, "Review of classifier combination methods," 2008.

[88] G.Webb, "Multiboosting: A technique for combining boosting and wagging," 2000.

[89] Y.S.Huang and C.Y.Suen, "A method of combining multiple experts for the recognition of unconstrained handwritten numerals," vol. 17, no. 1, January 1995.

[90] Fabio Roli, Lorenzo Bruzzone Giorgio Giacinto, "Combination of neural and statistical algorithms for supervised classification of remote sensing images," 2000.

[91] Imran Naseem, "Combining classifiers using the dempster shafer theory of evidence," dhahran, saudi arabia, january 2005.

[92] David H.Wolpert, "Stacked generalization," Los Alamos,NM,.

[93] Breiman L., "Stacked regressions machine learning," vol. 24, 1996.

[94] & D. Wolpert Symth P., "Stacked density estimation," 1997.

[95] Ian H. Witten Kai Ming ting, "Issues in stacked generalization," vol. 10, 1999.

[96] F.Alkoot and J.Kittler, "Experimental evaluation of expet fusion strategies," vol. 20, 1999.

[97] Prasanna ballal, "Dempster Shafer theory," 2004.

[98] Kari Sentz, "Combination of evidence in dempster shafer theory," Binghamton University,.

[99] Yager R.R, "On the Dempster Shafer Framework and new combination rules," 1987.

[100] D.Flangan, J.Zhang, and B.Rosenblatt J.Gotman, "Automatic seizure detection in the newborn: Methods and initial evaluation," 1997.

[101] R.J. Vermeulen, R.L.Strijers, W.P.Fetter and C.J.Stam J.Altenburg, "Seizure detection in the neonatal EEG with synchronization likelihood".

[102] Maromi Nei, "Cardiac Effects of Seizures," 2009.

[103] E.M.Vriens, F.S.S.Leijten, J.J.Spijkstra, A.R.J.Girbes, A.C.Van Huffelen, and C.J.Stam A.J.C.Slooter, "Seizure detection in adult ICU patients based on changes

in EEG synchronization likelihood," 2006.

[104] http://www.illustrationsource.com/stock/image/281952/illustration-of-the-brain-lateral-view-shown-within-an-outline-of-a-head/,.

[105] http://nursingcrib.com/nursing-notes-reviewer/seizure-disorder/,.

Curriculum Vitae

Name: Mohammed Abdul Azeem Siddiqui

Birth: 12th December, 1987, Hyderabad, INDIA.

Nationality : INDIAN

Education:

BACHELOR OF ENGINEERING (B.E)

Electronics and Communication,

Osmania University,

Hyderabad, INDIA.

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

MINOR : Signal Processing & Communication

Department of Electrical Engineering

King Fahd University of Petroleum & Minerals

Dhahran, K.S.A.

Email id: abdulazeem@kfupm.edu.sa, azeem012001@gmail.com,

Present Address:

P.O.Box # 1622,

King Fahd University of Petroleum & Minerals,

Dhahran -31261, KSA

Ph. No : +966-536611370

Permanent Address:

17-1—12/A, Flat no# 304, ALM apartments,

Old Santoshnagar,

Hyderabad, INDIA -500059

Ph No : +91-9948531867