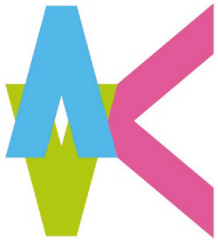
Classification of Epileptic Seizure EEG Signals in Time Frequency Domain - Focusing on Root Mean Square based Feature Extraction

A Thesis
submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

by

Arindam Gajendra Mahapatra
(Student Number: 15899028)



Kyutech
Kyushu Institute of Technology

Graduate School of Life Science and Systems Engineering
Kyushu Institute of Technology
Japan
June, 2018

Classification of Epileptic Seizure EEG Signals in Time Frequency Domain - Focusing on Root Mean Square based Feature Extraction

by

Arindam Gajendra Mahapatra

Submitted to the Graduate School of Life Science and Systems Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Author .....
Graduate School of Life Science and Systems Engineering
June, 2018

Certified by .....
Dr. Keiichi Horio
Associate Professor
Thesis Supervisor

Accepted by .....
Dr. Kyohisa Natsume
Chairman, Department Committee on Graduate Thesis

at the

Kyushu Institute of Technology

June, 2018

This doctoral thesis has been examined by a Committee of the
Graduate School of Life Science and Systems Engineering:

Dr. Kiyohisa Natsume .....
Chairman, Thesis Committee
Professor of Department of Human Intelligence System,
Kyushu Institute of Technology.

Dr. Hideaki Kawano .....
External Member, Thesis Committee
Associate Professor of Depart. of Electrical Engg. and Electronics,
Kyushu Institute of Technology.

Dr. Hiroaki Wagatsuma .....
Member, Thesis Committee
Associate Professor of Department of Human Intelligence System,
Kyushu Institute of Technology.

Dr. Keiichi Horio .....
Thesis Supervisor, Thesis Committee
Associate Professor of Department of Human Intelligence System,
Kyushu Institute of Technology.

Classification of Epileptic Seizure EEG Signals in Time Frequency Domain - Focusing on Root Mean Square based Feature Extraction

by

Arindam Gajendra Mahapatra

Submitted to the Graduate School of Life Science and Systems Engineering
on June, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Epilepsy affects over 50 million people on an average yearly world wide. Epileptic Seizure is a generalised term which has broad classification depending on the reasons behind its occurrence. Parvez et al. when applied feature instantaneous bandwidth B^2_{AM} and time averaged bandwidth B^2_{FM} for classification of interictal and ictal on Freiburg data base, the result dipped low to 77.90% for frontal lobe whereas it was 80.20% for temporal lobe compare to the 98.50% of classification accuracy achieved on Bonn dataset with same feature for classification of ictal against interictal. We found reasons behind such low results are, first Parvez et al. has used first IMF of EMD for feature computation which mostly noised induce. Secondly, they used same kernel parameters of SVM as Bajaj et al. which they must have optimised with different dataset. But the most important reason we found is that two signals s1 and s2 can have same instantaneous bandwidth. Therefore, the motivation of the dissertation is to address the drawback of feature instantaneous bandwidth by new feature with objective of achieving comparable classification accuracy. In this work, we have classified ictal from healthy nonseizure interictal successfully first by using RMS frequency and another feature from Hilbert marginal spectrum then with its parameters ratio. RMS frequency is the square root of sum of square bandwidth and square of center frequency. Its contributing parameters ratio is ratio of center frequency square to square bandwidth. We have also used dominant frequency and its parameters ratio for the same purpose. Dominant frequency have same physical relevance as RMS frequency but different by definition, i.e. square root of sum of square of instantaneous bandwidth and square of instantaneous frequency. Third feature that we have used is by exploiting the equivalence of RMS frequency and dominant frequency (DF) to define root mean instantaneous frequency square (RMIFS) as square root of sum of time averaged bandwidth square and center frequency square. These features are average measures which shows good discrimination power in classifying ictal from interictal using SVM. These features, f_r and f_d also have an advantage of overcoming the draw-

back of square bandwidth and instantaneous bandwidth. RMS frequency that we have used in this work is different from generic root mean square analysis. We have used an adaptive thresholding algorithm to address the issue of false positive. It was able to increase the specificity by average of 5.9% on average consequently increasing the accuracy. Then we have applied morphological component analysis (MCA) with the fractional contribution of dominant frequency and other rest of the features like bandwidth parameter's contribution and RMIFS frequency and its parameters and their ratio. With the results from proposed features, we validated our claim to overcome the drawback of instantaneous bandwidth and square bandwidth.

Keyword

Electroencephalography (EEG), Empirical Mode Decomposition (EMD), Hilbert Transform (HT), Root Mean Square (RMS), Dominant Frequency (DF), Root Mean Instantaneous Frequency (RMIFS), Morphological Component Analysis (MCA), Undecimated Wavelet Decomposition (UDWT), Local Discrete Cosine Transform (LDCT), Dirac, Support Vector Machine (SVM).

In the memory of my father who could not be here to see this day.

Late. Debkumar Gajendra Mahapatra

Acknowledgments

First of all, I praise God for providing me this opportunity and granting me the capability to proceed successfully. I write this note for acknowledgment in my thesis. It has been a period of intense learning for me in scientific arena as well as personal level. I would like to imitate on the people who have supported and encouraged me with suggestion in various direction and helped me so much throughout this period. First of all, I offer my sincerest gratitude to my supervisor, Associate Professor, Dr. Keiichi Horio, who has supported me throughout my thesis with his knowledge whilst allowing me to work in my own interest. I would like to thanks the member of PhD committee Professor, Dr. Kiyohisa Natsume, Associate Professor, Dr. Hiroaki Wagatsuma for their excellent suggestions and detail review during the thesis evaluation. In addition, I would like to thank my friend Dr. Balbir Singh for the valuable help. I cherish all the beautiful evening spent discussing with him. I thank all my lab members for their support, help and bearing my inapt Japanese communication skill all these years.

Last but not the least, I would like to thank my mother Smt. Snigdha Gajendra Mahapatra and my sister Mrs. Anuradha Maiti for their uncountably infinite patience and support all these time upon me. Thank you so much, you were always there for me. Thank you very much everyone who directly or indirectly supported and helped me completing my research.

Arindam Gajendra Mahapatra
Kyushu Institute of Technology
June, 2018

There is no expedient to which man will not resort to evade the real labor of thinking.

- Joshua Reynolds

The person who reads too much and uses his brain too little will fall into lazy habits of thinking.

- Albert Einstein

If you want to find the secrets of the universe, think in terms of energy, frequency and vibration.

- Nikola Tesla

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Electroencephalogram (EEG) is first hand tool to diagnosis epileptic seizures. Epilepsy affects over 50 million people world wide. Epileptic Seizure is a generalised term which has broad classification depending on the reasons behind its occurrence [1]. Epileptic seizure or ictal can be considered as hyperactivity of neural network which disrupts normal brain functioning from few seconds to several minutes. Epilepsy can be characterised by the predisposition to create an epileptic seizure. The word seizure having a Greek origin where it means *to take hold* [2]. Epileptic seizures have a clear start, but often termination of it is less evident than its onset. Interictal is period between two consecutive ictal or seizure. Seizure morphology depends on its onset in the brain its propagation, the maturity of the subject, any confounding disease, sleep cycle and variety of factors. The seizure can affects ones sensory, motor, emotional state, memory, behaviour and cognition too [2].

1.1 Background Study

In the past, for epileptic EEG classification, Fourier based methods were in use [3] [4] [5], but these methods have fixed basis functions. Fourier analysis requires data to be stationary and linear. It needs additional harmonic components to define non-stationary components. Short time Fourier transform (STFT) based time frequency

methods had also been used [6]. In STFT, there exist trade off between frequency and time resolution depending on the window size [7]. Wavelet are used for filtering purpose of EEG to decompose it [8]. Wavelet gives component with multiple levels of resolution [9]. Wavelet based analysis [7] [10] [11] [12] [13] [14] has also shown good results in classifying seizures because it can identify changes with respect to time, localised in high frequency range as epileptic seizures are. Guo et al. [15] have employed discrete wavelet transform to pre-analyse the EEG signals. Subasi et al. [16] has used DWT to decompose the EEG then by using principal component analysis (PCA), independent component analysis (ICA) and linear discriminant analysis to reduce the dimensionality of the data before feeding it to support vector machine (SVM).

In conjunction with each method, feature extraction remains one of the important part of EEG classification process, influencing their results. The results are as good as the selection and combination of the features. Line length based algorithms which give change in amplitude and frequency been reported [17]. Features like correlation dimension [18], fractal dimensions are applied previously [19]. Liang in his research work [20] had extracted approximate entropy (ApEn) along with other spectral features. These features were used with auto regressive model and principal component analysis (PCA), was able to quantify irregularities of signals. PCA been also used with sliding window based features in [21]. Peak value, equivalent width and mean square abscissa features are derived from cross-correlation of signals and power spectral density was applied in [22]. Energy and curve length of the signal feature were extracted using genetic programming in [15]. Tang et al. [23] had used median Teager energy with SVM assembly for seizure detection. Tempko et al. [24] had used total of 55 features, total power, peak frequency, normalised power, wavelet energy, spectral edge frequency are few of the spectral features. In time domain the features were curve length, number of maxima minima, autoregressive modelling error, skewness, kurtosis, non-linear energy, Hjorth parameters, zero crossing, root mean squared amplitude, nonlinear energy. Information theory based features were Shanon entropy, singular value decomposition entropy, spectral entropy, Fisher information

to name a few. SVM is used for classifying different type of ictal events in [23]. Permutation entropy feature was applied in [25] to classify ictal data from non-seizure and interictal data for creation of seizure detection system. Strength and degree of horizontal visibility graph (HVG) features, a type of complex network were also used by Zhu et al. [26] for ictal classification purpose. Higher order spectra based analysis also been done [27]. Early work using artificial neural network (ANN) based work with linear feature like amplitude, slope , curvature can be found in [28]. Prediction of preictal state to ictal state using Markov model was presented in [29]. Hassan et al. had employed spectral features to be used with various classifiers such as Parzen probabilistic neural network, discriminant analysis (DA), naive Bayes, artificial neural network (ANN), K-nearest neighbour (kNN), extreme learning machine (ELM), SVM, restricted Boltzman machine (RBM), bootstrap aggregating (Bagging), LS-SVM, random forest (RF) and adaptive boosting (AdaBoost). Boosting has shown best result among all classifiers [30].

Empirical mode decomposition (EMD), introduced by Huang et al. [31] has been used successfully to classify ictal EEG. In conjunction with EMD, parameters like weighted frequency [32], the coefficient of variation, fluctuation index [33], mean, standard deviation, variance, skew, centroid [34] [35] were employed for classification. Bandwidth amplitude modulation and bandwidth frequency modulation features are used for discriminating seizure from non-seizure using least square support vector machine (LS-SVM) in [36] along with EMD. Phase space representation was utilised to discriminate interictal from ictal by Sharma et al. [37]. RMS frequency was used along with another feature based on amplitude from Hilbert spectrum for the classification of seizure and non-seizure by us [38].

In epilepsy, the commonly observed behaviour or morphology is spike train, sharp waves. The sudden transient burst of spikes and high frequency oscillation in interictal recording are also used for the localisation of the epileptic seizures. Both, disparity in background activity and EEG paroxysms make the automated analysis complicated. Artefacts in filtered data can give rise to false positive [28-30]. Recently, signal decomposition by focusing morphological components are getting highlighted due to the

applicability to nonlinear and non-stationary signal properties [39, 40, 41].The mixing of sources cause the EEG signal to be non-linear and non-stationary in nature. Due to this, separation of sources from desired mixed signal become more difficult in time or frequency domain. MCA uses the linear combination of coefficients similar to independent component analysis (ICA). PCA and ICA [42] are popular methods used for separation of sources or removal of artefacts. Both methods works on a statistical approach, aim to find the linear projection of the signals, i.e. statistically independent [42]]. The subspace projection is used to extract EEG components on time/space basis. PCA is a sophisticated method to reduce the artefacts and specifies principal components (PC) to reconstruct overall data structure and to remove the components with small amplitudes and irregular changes. It is very difficult to specify remaining PCs to represent such signal. To identify PC, requires the prior knowledge of the artefacts [43]. In ICA, different estimation procedure such as mutual information minimization, maximization of non-Gaussianity, maximization of likelihood, SOBI, Fastica are used for separation. Since ICA is based on the measure of statistical independence, the noise of the input is amplified by ICA and it makes the detection of the signal components difficult due to Gaussian noise spread over the component in an undesired way [44]. ICA generates spikes and bumps, if the sample size is not sufficient [45]. ICA is a multichannel source separation techniques and doesn't work on single channel unlike MCA which can work perfectly with single channel [45]. There are other methods also presented based on sparse classification [46, 47]. Now days, tensor based works are also coming forward on EEG using some deep learning framework. The classification results are fascinating but often miss the reasoning behind it and focussing on hyper-parameters tuning of deep network [48].

1.2 Motivation and Objective of Dissertation

Parvez et al. [49] when applied feature averaged instantaneous bandwidth square B_{AM}^2 and time averaged bandwidth square B_{FM}^2 for classification of interictal and ictal on Freiburg data base, the result dipped low to 77.90% for frontal lobe whereas

it was 80.20% for temporal lobe compare to the 98.50% of classification accuracy achieved on Bonn dataset with same feature for classification of ictal against interictal. We found reasons behind such low results are, first Parvez et al. has used first IMF of EMD for feature computation which mostly noised induce. Secondly, they used same kernel parameters of SVM as Bajaj et al which they must have optimised with different dataset. But the most important reason we found is that two signals s1 and s2 can have same instantaneous bandwidth. Therefore, the motivation of the dissertation is to address the drawback of feature instantaneous bandwidth by new feature with objective of achieving comparable classification accuracy. In this work, we have classified ictal and interictal successfully first by using RMS frequency and another feature from Hilbert marginal spectrum then with its parameters ratio. RMS frequency is the square root of sum of square bandwidth and square of center frequency. Its contributing parameters ratio is ratio of center frequency square to square bandwidth. We have also used dominant frequency and its parameters ratio for the same purpose. Dominant frequency have same physical relevance as RMS frequency but different by definition i.e. square root of sum of square of instantaneous bandwidth and square of instantaneous frequency. Third feature that we have used is by exploiting the equivalence of RMS frequency and dominant frequency (DF) to define root mean instantaneous frequency square (RMIFS) as square root of sum of time averaged bandwidth square and center frequency square. These features are average measures which show good discrimination power in classifying ictal from interictal using SVM. These features, f_r and f_d also have an advantage of overcoming the drawback of square bandwidth and square instantaneous bandwidth. RMS frequency that we have used in this work is different from generic root mean square analysis as in [50], [51]. We have used an adaptive thresholding algorithm to address the issue of false positive. It was able to increase the specificity by the average of 5.9% on average consequently increasing the accuracy. Then we have applied morphological component analysis (MCA) with the fractional contribution of dominant frequency and other rest of features like bandwidth parameter's contribution and RMIFS frequency and its parameters and their ratio.

1.3 Methodology and Organisation of Dissertation

In this work, the methodology we have followed is first decomposed the EEG from Bonn dataset using first with EMD then in second part of the work, by using MCA. Apply Hilbert transform (HT) over the given output to take the real valued EEG time series into complex domain. This enables us to represent it into analytic form which helps in computing instantaneous frequency and amplitude component of the output. Compute the features and normalised it using mean and deviation before feeding it to SVM for classification of ictal from interictal and healthy non seizure. Evaluate the output of SVM with standard statistical parameters and compare it other works. This dissertation is organised as EMD and HT are described in second chapter, third chapter contain initial result with RMS frequency and proposal of adaptive thresholding method. Followed by the fourth chapter where we have improved the initial results and proposed new feature root mean instantaneous frequency based on relation between dominant frequency and RMS frequency. The fifth chapter contain the application of MCA with rest of the features proposed in third and fourth chapter. Concluding with a comparison of EMD and MCA based on EEG classification results of epileptic seizures on Bonn dataset.

1.4 Summary

The ability to discriminate between the ictal from non-seizure and interictal EEGs of epileptic patients is important for practical applications like seizure detection, prediction, warning systems or closed-loop seizure control systems [20], [52]. To create seizure detection system, we need features to determine presence of seizure from interictal data [53]. The first step is to have features showing good classification accuracy with high sensitivity for ictal and interictal data. The standard process for classification of EEG that is been followed is divided in three parts mainly. The first part is

decomposition or filtering process which is used as preprocessing but optional. These step is important as the EEG we get is summation of neural, sub neural network activity or culmination of thousand of neurones synchronous activity. Hence decomposing these activity into oscillatory modes depicting particular brain state is important. Many researchers instead of using any decomposition method, apply band pass based filter. But using bandpass filter may add some spurious harmonics in the oscillations. Second comes feature extraction, it is very important component of any method as feature extracted directly affects the classification results. Features are extracted mainly as spectral, time frequency, energy, entropy and some non linear parameters. In this research work, we are focussing on time frequency based feature extraction. Third comes, supervised classification algorithm mainly but other machine learning techniques are also been in use to distinguish EEGs. Standard methodology is presented in Figure 1-1. In this work, the focus is on feature extraction with belief good feature lead to excellent classification result. We have also address the problem of identical instantaneous bandwidth square and bandwidth square while classifying ictal, non ictal EEGs.

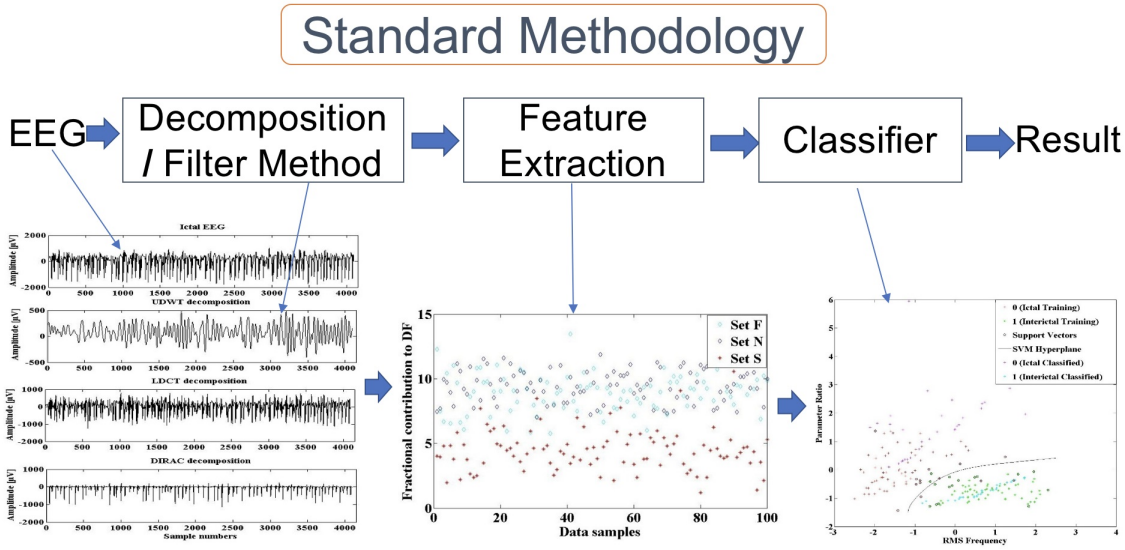


Figure 1-1: Methodology followed in this work.

Chapter 2

Empirical Mode Decomposition

2.1 Introduction

In this chapter, we have discussed the empirical mode decomposition (EMD) as decomposition method applied in this work. Empirical decomposition method is useful in handling the non-linear non stationary EEG signal. Hilbert transform when applied over the amplitude modulated and frequency modulated output of EMD, called intrinsic mode functions (IMFs) gives instantaneous frequencies as the function of time.

2.2 Empirical Mode Decomposition

EMD [31] analyse a signal adaptively without imposing any predefined model, completely dependent on the data. It decomposes the EEG signal into the number of oscillatory modes called intrinsic mode function (IMF).
EMD accept a decomposition as an IMF on two conditions :

1. The difference between the number of zero crossing and number of extrema must be either 0 or at most 1.

2. The mean calculated at any instant of time from the upper and lower envelope must result into zero.

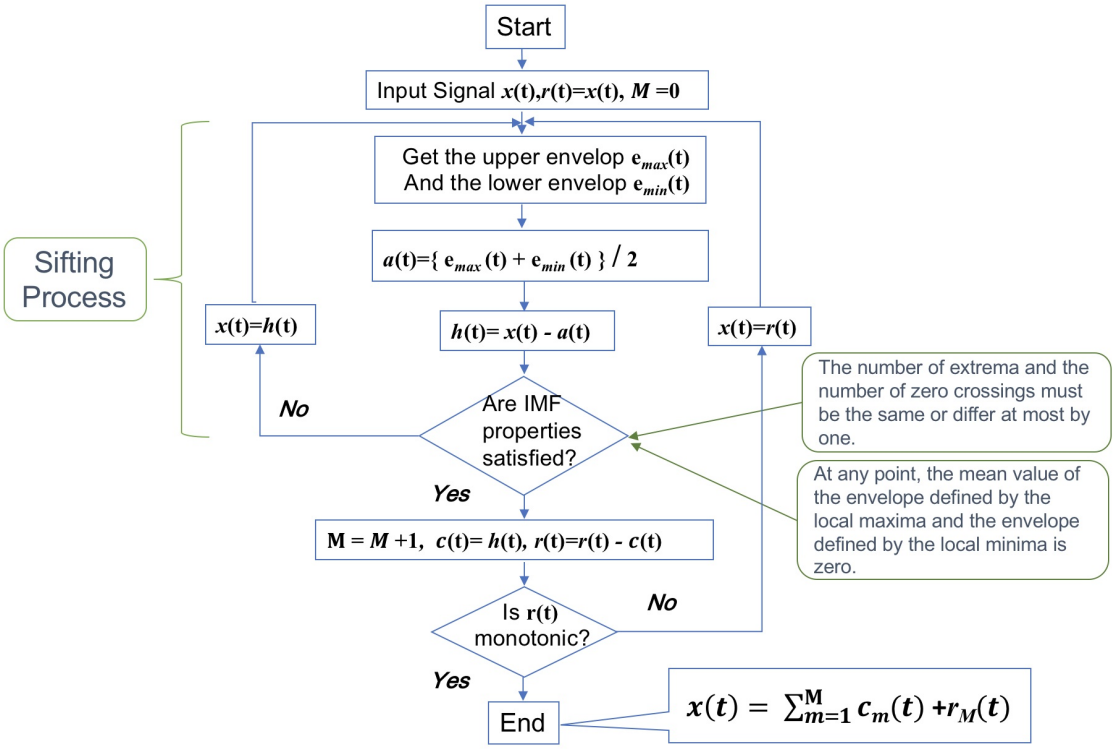


Figure 2-1: EMD block diagram.

Algorithm is depicted in Figure 2-1

Recursively apply the whole presented process called sifting process until residue $r(t)$ conform to be monotonic. In the end, the original signal can be given by the summing all the IMF decompositions and the monotonic residue as represented in equation (2.1). Figure 2-2 depicts EMD decomposition of interictal EEG signal and Figure 2-3 presents the decomposition of ictal EEG. We have presented first eight decompositions or IMFs along with the residue in the figures.

$$x(t) = \sum_{m=1}^{M} c_m(t) + r_M(t), \tag{2.1}$$

where M represent total number of IMFs and $c_m(t)$ represent the m^{th} IMF. $r_M(t)$ is the monotonic residue at the end.

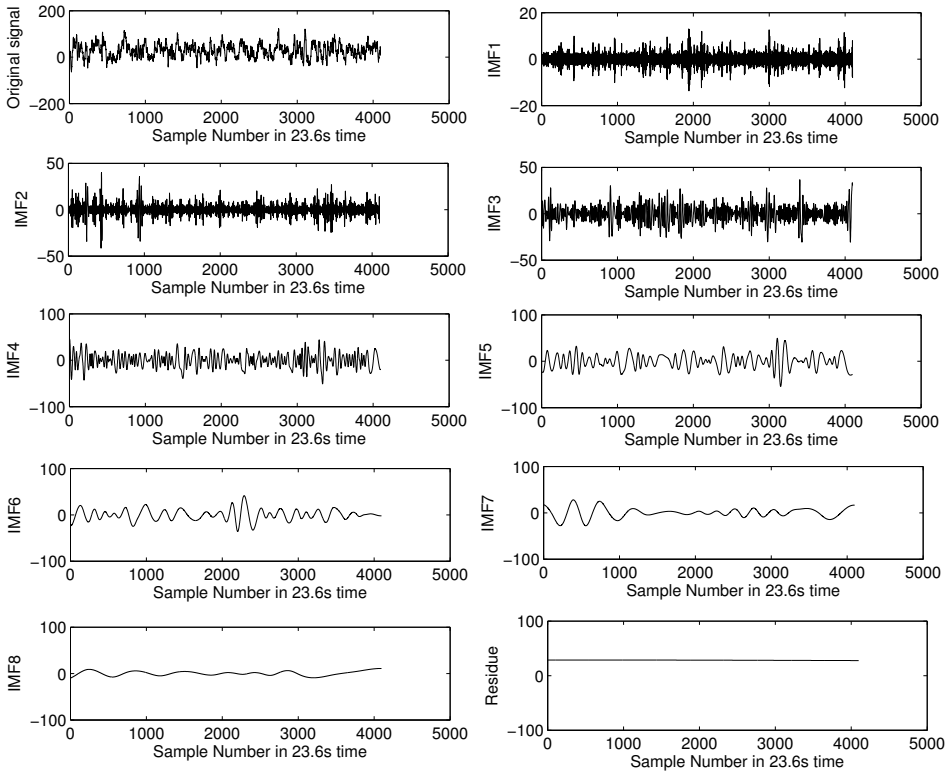


Figure 2-2: EMD decomposition of interictal EEG signal.

2.3 Hilbert Transform

Hilbert transform was applied on the IMF produced by the EMD. It takes the real valued IMF $c(t)$ to complex time frequency domain by representing it in analytic form as

$$s(t) = c(t) + jc_H(t), \tag{2.2}$$

where applying Hilbert transform on $c(t)$ gives $c_H(t)$ defined as $c_H(t) = c(t) * \frac{1}{\pi t}$, $*$ represents convolution and if we take inverse of Fourier transform of $\frac{1}{\pi t}$, it will give $-jsgn(\omega)$ (where ω represents signal's phase) signifying the difference of phase of $\frac{\pi}{2}$ between positive frequency and negative frequency. We ignore the negative frequency represented by imaginary part. By utilizing hermitian symmetry, we work only with the real part. So, we take the real valued IMF to the time frequency domain by projecting it on the real axis of the complex domain. Equation (2.2) can be represented

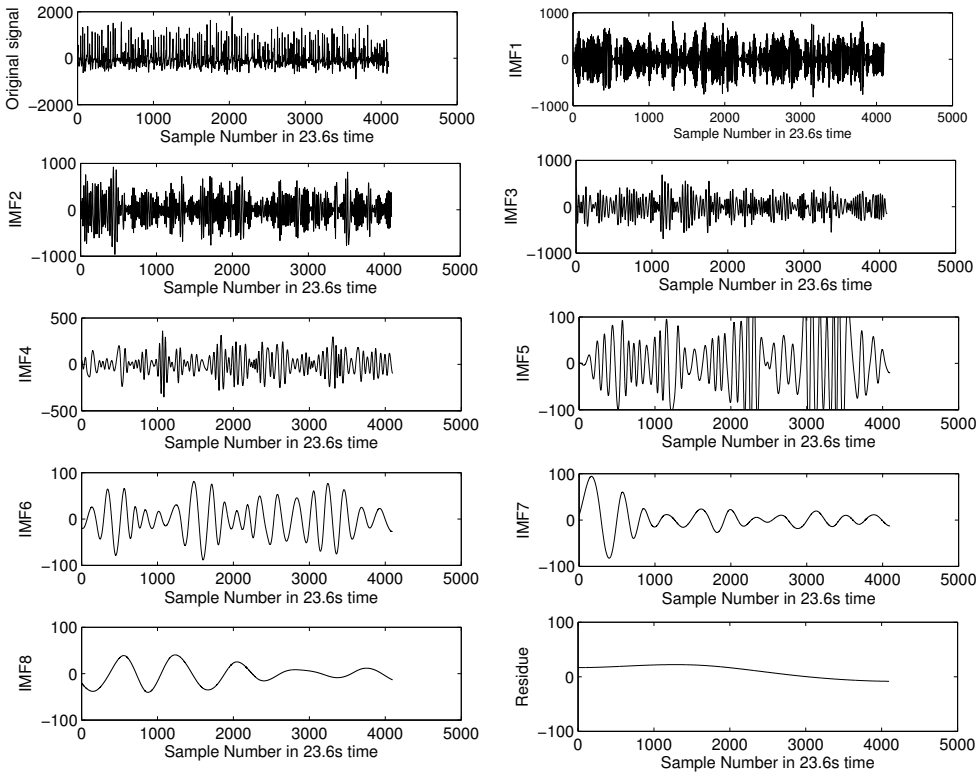


Figure 2-3: EMD decomposition of ictal EEG signal.

as in [31]

$$s(t) = a(t)e^{j\varphi(t)}. \tag{2.3}$$

Instantaneous phase $\varphi(t)$ and amplitude $a(t)$ can be given by

$$\varphi(t) = \arctan\left[\frac{c_H(t)}{c(t)}\right]. \tag{2.4}$$

$$a(t) = \sqrt{c^2(t) + c_H^2(t)}, \tag{2.5}$$

Instantaneous frequency is defined as derivative of instantaneous phase as in [54]

$$\omega(t) = \varphi'(t). \tag{2.6}$$

Prime represents differentiation.

EEG signals are non-stationary in nature [55]. EMD adaptively decomposes EEG exhibiting intrawave frequency modulations into set of IMFs expressing from highest oscillatory modes to lowest oscillatory mode containing AM and FM components. EMD decomposes a signal by considering fast oscillations called the local detail superimposed upon slow oscillation called as the local trend. By putting condition for IMF to be zero mean it ensure that all maxima have to be positive and all the minima have to be negative. Recursively extracting the fast oscillation from slow oscillation until we get monotonic residue is call sifting process. IMFs are defined in this way so that it can exhibit locality in time [56], [57].

After applying Hilbert transform over IMF which bring the real value decomposition to the complex plane. This enables to represent the IMF in analytic form, taking derivative of instantaneous phase gives a single value function of time i.e. at any given instant of time there is only one frequency, therefore mono component [31]. In absence of any clear definition of mono-component signal, data to be narrowband was put as an restriction, i.e. number of extrema and zero crossing to be equal i.e. second condition on decomposition to be called an IMF. To get more meaningful instantaneous frequency, restriction has been put on it to be positive frequency only which we get when the function is symmetric with respect to zero mean level. For these reasons, the two conditions are imposed over decomposition to be called an IMF.

2.4 Summary

EMD effectively decompose EEG signal adaptively and hence indefinite number of decompositions but there are many problem associated with EMD. The major problem is computational overhead pose by EMD before feature extraction as there is indefinite number of decomposition produced by EMD. Using Hilbert transform in projecting the decomposition on the real axis of complex domain helps avoiding working with negative frequency. Secondly, due to Hermitian symmetry no information is lost in this transformation. Real signal has symmetric spectrum which makes aver-

age quantities zero. Therefore, taking real signal to complex domain and working on positive frequency solves this problem [62].

Chapter 3

Root Mean Square Frequency

3.1 Introduction

After applying EMD, time frequency based feature root mean square and along with feature based on Hilbert marginal spectrum being calculated after using Hilbert transform. RMS frequency, we believe can overcome the drawback of square bandwidth which can be same for two signals at a time simultaneously and at the same time can be used as average measure for classifying epileptic seizure.

3.2 Features Computation

Mandel [58] and Populis [59] defined root mean square (RMS) frequency f_r. RMS frequency was mentioned by Barnes et al. [60] in their seismic application. RMS frequency is the square root of sum of squared center frequency and square bandwidth.

This can be shown from square bandwidth as follows from [61], [62].

$$\sigma_\omega^2 = \int (\omega - < \omega >)^2 |S(\omega)|^2 d\omega, \tag{3.1a}$$

$$= \int (\omega^2 + < \omega >^2 - 2\omega < \omega >) |S(\omega)|^2 d\omega, \tag{3.1b}$$

$$= \int (\omega^2) |S(\omega)|^2 d\omega + \int < \omega >^2 |S(\omega)|^2 d\omega - \int (2\omega < \omega >) |S(\omega)|^2 d\omega, \tag{3.1c}$$

$$= < \omega^2 > + < \omega >^2 \int |S(\omega)|^2 d\omega - 2 < \omega > \int (\omega) |S(\omega)|^2 d\omega, \tag{3.1d}$$

$$\sigma_\omega^2 = < \omega^2 >_S - < \omega >_S^2 . \tag{3.1e}$$

In this work, the square bandwidth σ_ω^2 is also represented by B^2 and all integrals are between time interval $[t_1, t_2]$. Rearranging the equation (3.1e) gives us root mean square frequency as given by [60].

$$< \omega^2 >_S = < \omega >_S^2 + \sigma_\omega^2, \tag{3.2a}$$

$$f_r = \sqrt{< \omega >_S^2 + B^2}. \tag{3.2b}$$

where $< \omega^2 >_S$ is mean square frequency and $< \bullet >_S$ stands for spectral domain (ω).

$$< \omega^2 >_S = \int \omega^2 |S(\omega)|^2 d\omega. \tag{3.3}$$

And $< \omega >_S$ represents spectral average or center frequency as in [61].

$$< \omega >_S = \int \omega |S(\omega)|^2 d\omega. \tag{3.4}$$

Time average $< \bullet >_T$ of the instantaneous frequency

$$< \omega >_T = \int \varphi'(t) |s(t)|^2 dt. \tag{3.5}$$

The $S(\omega)$ is the Fourier transform of the signal $s(t)$.

$$S(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-j\omega t} s(t) dt. \tag{3.6}$$

Applying hermitian time frequency operator $\left(\frac{1}{j}\frac{d}{dt}\right)$ to equation (3.4). Center frequency can be written as in [62]

$$< \omega >_S = \int \omega |S(\omega)|^2 d\omega, \tag{3.7a}$$

$$= \int s^*(t) \frac{1}{j} \frac{d}{dt} s(t) dt, \tag{3.7b}$$

$$= \int s^*(t) \left(\frac{1}{j} \frac{d}{dt} s(t) dt \right), \tag{3.7c}$$

$$= \int s^*(t) \left(\frac{1}{j} \frac{d}{dt} a(t) e^{j\varphi(t)} dt \right), \tag{3.7d}$$

$$= \int s^*(t) \left(\frac{1}{j} \Big\{ a'(t) e^{j\varphi(t)} + j\varphi'(t) a(t) e^{j\varphi(t)} \Big\} \right) dt, \tag{3.7e}$$

$$= \int s^*(t) \left(\frac{1}{j} \Big\{ a'(t) + j\varphi'(t) a(t) \Big\} e^{j\varphi(t)} \right) dt, \tag{3.7f}$$

$$= \int s^*(t) \left(\frac{1}{j} \Big\{ \frac{a'(t)}{a(t)} + j\varphi'(t) \Big\} a(t) e^{j\varphi(t)} \right) dt, \tag{3.7g}$$

$$= \int s^*(t) \left(\Big\{ \frac{1}{j} \frac{a'(t)}{a(t)} + \varphi'(t) \Big\} s(t) \right) dt, \tag{3.7h}$$

$$= \int \left(\varphi'(t) + \frac{1}{j} \frac{a'(t)}{a(t)} \right) a^2(t) dt. \tag{3.7i}$$

Equation (3.7c) is simplified by substituting equation (2.3) in it. We have ignored the imaginary part as it will be zero, $s^*(t)$ is complex conjugate signal and $a^2(t)$ is density in time [61].

Therefore, center frequency is, as in [62]

$$< \omega >_S = \int \varphi'(t) a^2(t) dt = < \omega >_T . \tag{3.8}$$

If we compare equations (3.4) and (3.5) we see $< \omega >_S = < \omega >_T$ as in [63]. Square bandwidth can be similarly expressed as in [61] [62]

$$\sigma_\omega^2 = \int (\omega - <\omega>)^2 |S(\omega)|^2 d\omega, \tag{3.9a}$$

$$= \int s^*(t) \left(\frac{1}{j}\frac{d}{dt} - <\omega> \right)^2 s(t) dt, \tag{3.9b}$$

$$= \int \left| \left(\frac{1}{j}\frac{d}{dt} - <\omega> \right) s(t) \right|^2 dt, \tag{3.9c}$$

$$= \int \left| \frac{1}{j}\frac{a'(t)}{a(t)} + \varphi'(t) - <\omega> \right|^2 a^2(t) dt, \tag{3.9d}$$

$$= \int \left(\frac{a'(t)}{a(t)} \right)^2 a^2(t) dt + \int (\varphi'(t) - <\omega>)^2 a^2(t) dt, \tag{3.9e}$$

$$= \int a'^2(t) dt + \int (\varphi'(t) - <\omega>)^2 a^2(t) dt, \tag{3.9f}$$

$$B^2 = B_{AM}^2 + B_{FM}^2, \tag{3.9g}$$

$$B = \sqrt{B_{AM}^2 + B_{FM}^2}. \tag{3.9h}$$

Above equation (3.9f) signifies the contribution of both amplitude modulation (AM) and frequency modulation (FM) to square bandwidth where $a'(t)$ is derivative or change in amplitude. These parameters are independent of each other as B_{AM}^2 is based on amplitude and B_{FM}^2 is dependent on phase function $\varphi(t)$ which is an advantage. B_{FM}^2 will be small when instantaneous frequencies are close to center frequency, i.e. small frequency modulation. [61].

$$B_{AM}^2 = \int a'^2(t) dt, \tag{3.10}$$

$$B_{FM}^2 = \int (\varphi'(t) - <\omega>)^2 a^2(t) dt. \tag{3.11}$$

Fractional contribution to root mean bandwidth and ratio of AM and FM contribution as given by [61]

$$\tau_{AM} = \frac{B_{AM}^2}{B}, \tag{3.12}$$

$$\tau_{FM} = \frac{B_{FM}^2}{B}, \tag{3.13}$$

$$\tau_B = \frac{B_{AM}^2}{B_{FM}^2}. \tag{3.14}$$

Now, if we take equation (3.9e) and rewrite it as

$$\sigma_\omega^2 = \int \left\{ \left(\frac{a'(t)}{a(t)} \right)^2 + (\varphi'(t) - < \omega >)^2 \right\} a^2(t) dt, \tag{3.15a}$$

$$\sigma_\omega^2 = \int \sigma_\omega^2(t) a^2(t) dt + \sigma_T^2. \tag{3.15b}$$

where σ_ω^2 is global standard deviation or spectral square bandwidth and $\sigma_\omega^2(t)$ is square local standard deviation i.e. square instantaneous bandwidth. Lastly σ_T^2, the time averaged deviation of instantaneous frequency around global mean or center frequency. Therefore, we can say spectral bandwidth square always be greater than the instantaneous frequency spread around center frequency by average instantaneous bandwidth square.

Root mean square frequency is given by [60] and from equation (3.2b)

$$f_r = \sqrt{< \omega >_S^2 + B^2}, \tag{3.16a}$$

$$f_r = \sqrt{< \omega >_T^2 + B_{AM}^2 + B_{FM}^2} \tag{3.16b}$$

Adding center frequency square to bandwidth square makes the feature RMS frequency unique as it is highly unlikely that squared center frequency and squared bandwidth be identical simultaneously for two signal at any instant of time.

Taking cue form [61], we can define fractional contribution and their ratio.

$$\gamma_{RMS_{AM}} = \frac{B_{AM}^2}{f_r}, \tag{3.17}$$

$$\gamma_{RMS_{FM}} = \frac{B_{FM}^2}{f_r}, \tag{3.18}$$

$$\gamma_{RMS_{<\omega>_T^2}} = \frac{< \omega >_T^2}{f_r}, \tag{3.19}$$

$$\Gamma_{RMS} = \frac{< \omega >_T^2}{B^2}. \tag{3.20}$$

If we take different contribution ratio, then we can define some more features as

$$\Gamma_{RMS} = \frac{B^2}{< \omega >_T^2}, \tag{3.21}$$

$$\Gamma_{RMS_{AM}} = \frac{B_{AM}^2}{< \omega >_T^2}, \tag{3.22}$$

$$\Gamma_{RMS_{FM}} = \frac{B_{FM}^2}{< \omega >_T^2}. \tag{3.23}$$

Second feature was calculated using Hilbert marginal spectrum. Marginal spectrum can be defined as integral of Hilbert Huang Transform (HHT) time frequency spectrum [31]

$$M(\omega) = \int_0^\infty H(\omega, t)dt. \tag{3.24}$$

which gives us total amplitude contribution from each frequency scale. We named it Amplitude Contribution of Frequency (ACF). We have used sum of maximum ten amplitude from the Hilbert marginal spectrum to create our second feature

$$ACF = \sum_{i=1}^{10} max|M(\omega)|. \tag{3.25}$$

3.3 SVM

Support vector machine (SVM) is machine learning technique, introduced by Vapnik [64] for solving binary classification problem. Discrimination between distinct classes is done by drawing a hyper plane by SVM by maximizing the distance/margin between the classes. The kernel functions used by us in this work are as follows :
Radial Basis function (RBF) kernel is represented by

$$G(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(\frac{||\boldsymbol{x_i} - \boldsymbol{x_j}||^2}{2\sigma^2}\right), \tag{3.26}$$

where σ is a positive number.

Polynomial Kernel can be represented by

$$G(\boldsymbol{x}_i, \boldsymbol{x}_j) = (1 + \boldsymbol{x}_i^T \boldsymbol{x}_j)^p, \tag{3.27}$$

where p is the degree of polynomial.

Linear Kernel

$$G(\boldsymbol{x}_i, \boldsymbol{x}_j) =< \phi(\boldsymbol{x}_i, \boldsymbol{x}_j) > . \tag{3.28}$$

3.4 Material and Simulation

3.4.1 Dataset

We have applied our method on EEG dataset [65] commonly known as Bonn dataset. The dataset comprises 5 subset F, N, O, S and Z. Each subset contains 100 signals segments of 23.6 seconds duration. Each signal segment consist of 4097 samples. The recording sampling frequency was 173.61 Hz. Subset Z, O were recorded extracranially with eye closed and with eyes open from healthy subjects having no seizure history. Subset F, N and S have signal segments from intracranial experiments. Subset F and N has interictal (period between two consecutive seizure) recording. Subset N is from epileptic zone and F is from hippocampal formation of the opposite hemisphere. Subset S contains ictal EEG recording. First, subset Z, O, F, N and S are used by us. Though subset F, N are results of intracranial experiments and subset Z, O were extracranially recorded yet classification been done to comparison purpose. We have created two combination of sets for classification. One with sets Z, O, F, S, second with Z, O against S.

The best classification average accuracy is obtained 97.72% in IMF3 with polynomial kernel function of SVM with default parameter. But to know the effectiveness, consistency of feature, we have used linear kernel and polynomial kernel function with default parameter and RBF kernel with scale set to 1.0 and box constraint also set to 1.0. The results are shown in Table 1 and Table 2 which prove that with these

features, the sets are linearly classifiable. The motto to use the default parameter is not only to show that feature can consistently discriminate the classes but also to keep the method simple by avoiding parameter optimisation to makes it more of practical use. It gives ample amount of proof through results that it overcomes the drawback of instantaneous bandwidth feature. We have used 70-30 ratio of training and test sets and generated the those set by picking up the samples randomly. Initially we have gone with 10 trials first then to be sure we have taken 100 trials for each kernel function. We have presented the best result attained through polynomial kernel function. We have compared our results with work of Li et al. [66] and Fu et al. [67] as both have used EMD with SVM as we have. This helps us to prove our claim of overcoming the drawback of instantaneous bandwidth by showing the effectiveness and consistency of the proposed feature.

3.5 Results and Discussion

EEG signals exhibit intrawave frequency modulation which can be viewed in two ways. One is fast oscillating signal superimposed over a slow oscillatory signal [68] or where frequency within the signal changes intermittently. Univariate EMD is data driven method for time frequency analysis of real valued signals. Based on the local characteristics it decomposes the EEG signals into finite number of IMFs which express from highest oscillatory mode to lowest one. Applying Hilbert transform over the IMF brings them to time frequency domain where we can express it in analytical form, enabling us to calculate instantaneous frequency. Instantaneous bandwidth gives us the information of the evolution of the signal [61] and is measure of spread of the frequency. Barnes [60] in his work observed that relative amplitude changes in large rates in narrow signals resulting large bandwidth which is similar to the observation expressed by Bajaj et al. [36] that the non-seizure signals have large instantaneous bandwidth than the seizure signal. Work by Bajaj et al. [36] shows excellent classification results with the bandwidth parameters on Bonn dataset. But when Parvez et al. [49] had applied the same for classifying ictal and interictal EEG

signal over Freiberg dataset, the results dipped very low. The reason we believe is, square bandwidth is the frequency spread i.e. a number and it doesn't show from where it belongs, where it has happened. Secondly Loughlin et al. [69] in his work has proved that two signals can have identical instantaneous bandwidth square and [62] shown bandwidth square can also be same. Another reason we believe that authors have used first IMF for the classification which is highly probable that it was induced with noise and artefact at the time experiment. Authors have also used the same parameter of the kernel function as used by Bajaj et al. [36] which they should have tuned accordingly again depending on the creation of random train set and test set. RMS frequency which is defined as square root of sum of squares of center frequency and square bandwidth can deal with the mention drawback of square bandwidth. Reason for our belief is that adding center frequency with the bandwidth square acts as a base which can identify where the spread has occurred and secondly it is highly improbable that center frequency and instantaneous frequency can be same for two signals at times. We have created and used another feature from the Hilbert marginal spectrum to supplement the RMS frequency. Using the results of Bajaj et al. [36] of Krushkal Wallis test and p-value as they have used same dataset and EMD code as done by us. We have decided to use first four IMF for generating the feature for classification as it shows statistical significance. RMS frequency and sum of maximum ten amplitude from Hilbert marginal spectrum were fed to SVM for classification and parameters for test performance are used as utilised in previous works, i.e. sensitivity (SEN), specificity (SPE) and Accuracy (Acc) [20], [36].

$$SEN = \frac{TP}{TP + FN} \times 100, \tag{3.29}$$

$$SPE = \frac{TN}{TN + FP} \times 100, \tag{3.30}$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \times 100, \tag{3.31}$$

where TP stands for true positive event i.e. detecting ictal correctly. FN stands

Table 3.1: Result of Classification of non seizure (Z, O, F, N) and seizure (S) of Bonn dataset.

EMD	Li et al. [66]		Kernel Function	Work results with 10 trials 70-30 ratio			
IMF	SEN[%]	SPE[%]		SEN[%]		SPE[%]	
				min-max	avg	min-max	avg
IMF1	94.75	98.00	Polynomial (p=5)	86.00-100.0	93.66	86.66-100.0	93.66
IMF2	93.25	97.50	Polynomial (p=5)	90.00-100.0	96.33	79.16-85.00	83.00
IMF3	79.00	95.25	Polynomial (p=3)	90.00-100.0	96.33	90.83-95.00	92.75
IMF4	72.75	86.75	Polynomial (p=3)	73.33-86.66	79.66	88.33-92.25	90.25

for false negative. FN signifies the failure to identify ictal which it has labelled as interictal. FP shows false positive. TN represents true negative event.

Table 3.2: The results with normal (Z,O) and ictal (S) EEG using EMD with polynomial kernel function (parameter p=3) of SVM on Bonn Dataset.

EMD	Li et al. [66]		Work results with 100 trials 70-30 ratio			
IMF	SEN[%]	SPE[%]	SEN[%]		SPE[%]	
			min-max	avg	min-max	avg
IMF1	93.25	96.90	93.33-100.0	98.06	93.33-100.0	98.06
IMF2	87.50	94.50	90.00-100.0	96.43	50.00-71.66	61.63
IMF3	85.25	94.40	90.00-100.0	96.43	95.00-100.0	98.36

The highest average classification accuracy was 97.72% (93.33-100.0%) seen in IMF3 using polynomial kernel with default parameter p=3 when applied on non-seizure set (Z, O) and seizure set (S). Fu et al. [67] has achieved maximum accuracy 99.125% (97.50-100%) with RBF kernel for theta wave by using only set Z and S. To compare with it, we have also used the same RBF Kernel function with parameter scale set to 1.0 and box constraint set to 1.0 over set Z and S. We have also used same 60-40 ratio of training set and testing set as Fu et al. [67] has used. We have attained 96.025 % of average accuracy. The best result we attained is with IMF3 where accuracy reached 100.0% (90.0-100.0%). Sensitivity was observed between 80.0-100.0% with maximum 100.0%. Specificity shown consistency of 100.0% (100.0-100.0%).

Taking 100 random trials is the reason for getting a wide range of result than Fu et al. [67] but the results confirm the consistency of the feature RMS frequency with an acceptable result.To get more affirmative about the feature, we have tested it for classifying ictal and interictal set. The Best result we got again with the polynomial kernel with default parameter p=3. The average accuracy are, in IMF1 it was 94.08% (89.16-97.50%), IMF2 showed 94.17% (90.83-97.50%) and IMF3 have 91.41% (85.00-95.83%) which is much better than what Parvez et al. observed 77.90% for frontal lobe and 80.20% for temporal lobe. As to compensate for using Bonn dataset instead of Freiberg dataset which Parvez et al. [49] has used, we have taken 100 random trials with 60% of feature for training and 40% of feature for testing as used by Bajaj et al. [36] in their experiment with instantaneous bandwidth parameter feature which confirms the effectiveness of RMS frequency.

3.5.1 Adaptive Thresholding

Another method based on threshold was also developed to counter drawback of B_{AM}^2. Classification of ictal EEG from interictal using bandwidth amplitude modulated (B_{AM}^2) and frequency modulated (B_{FM}^2) components over Support Vector Machine (SVM) shows average specificity as high as 96.76% but average sensitivity remain as the low as 50.43% resulting into low accuracy of 73.60%. Using $\frac{B_{AM}^2}{B_{FM}^2}$ as threshold to recognise the false negative from the output of SVM recovers the low sensitivity consequently accuracy. B_{AM}^2 and B_{FM}^2 values from interictal set are in lower range than the ictal signals and congregate at one place when mapped in two dimensions. This helps SVM to create hyperplane totally separating interictal. But with interictal (true negative) many ictal (true positive) features got conglomerated resulting into high false negative consequently low sensitivity and low accuracy. $\frac{B_{AM}^2}{B_{FM}^2}$ values of interictal are at higher range than ictal signals. It infers that in interictal, B_{AM}^2 is greater than B_{FM}^2 and dominates the bandwidth. Applying minimum value of $\frac{B_{AM}^2}{B_{FM}^2}$ corresponding to interictal training set of B_{AM}^2, B_{FM}^2 as threshold to identify the false negative in the corresponding $\frac{B_{AM}^2}{B_{FM}^2}$ to the output of SVM helps in recovering the low sensitivity and accuracy. Average results of 100 trials are shown

in Table 3.3. Accuracy of 99.46% is comparable with other recent works on ictal interictal classifications.

Table 3.3: Result of classification B^2_{AM} and B^2_{FM} with threshold $\frac{B^2_{AM}}{B^2_{FM}}$ on interictal (F, N) and ictal (S) of Bonn dataset for 100 trials.

		Without Threshold		With Threshod	
Kernel Function		IMF1 [%]	IMF2 [%]	IMF1 [%]	IMF2 [%]
RBF	SEN	57.66	71.73	100.0	100.0
	SPE	100.0	100.0	98.93	98.50
	Acc	**78.83**	**85.86**	**99.46**	**99.25**
Polynomial	SEN	61.36	77.73	100.0	100.0
	SPE	100.0	99.06	98.63	97.80
	Acc	**80.68**	**88.40**	**99.31**	**98.90**
Linear	SEN	56.50	59.99	100.0	100.0
	SPE	100.0	98.33	98.63	98.76
	Acc	**78.25**	**79.11**	**99.31**	**99.38**

3.6 Summary

The RMS frequency, we believe overcomes the drawback of the square bandwidth feature which may be identical for two signals as we have taken 10-100 random trials to be sure with the results. Though the range of the results are large yet the average results shows consistency with default parameter of basic kernel functions, i.e without any kind of parameter optimisation. We have got the best results with polynomial kernel consistently. We believed if we use better feature than sum of max amplitude of Hilbert marginal spectrum with signal being normalised, we expect to get more accurate results having physical relevance. We believe feature from Hilbert marginal

spectrum has pulled the result down as it is difficult to discriminate interictal and ictal set based on amplitude. And as we could not use any well known, well validated feature because then it will make it difficult for us to prove the effectiveness of feature RMS frequency. In fractional contribution, numerator was kept in square form to create more significant differences between the set without affecting the physical meaning of the feature, i.e. square standard deviation and centre frequency. Hence feature fractional contribution can also be defined in terms of standard deviations as

$$\gamma_{RMS_{AM}} = \frac{\sqrt{B_{AM}^2}}{f_r} = \frac{B_{AM}}{f_r}, \tag{3.32}$$

$$\gamma_{RMS_{FM}} = \frac{\sqrt{B_{FM}^2}}{f_r} = \frac{B_{FM}}{f_r}, \tag{3.33}$$

$$\Gamma_{RMS} = \frac{\sqrt{<\omega>_T^2}}{B^2} = \frac{<\omega>_T}{f_r}. \tag{3.34}$$

Secondly, if we look at the feature B^2 or mean square f_r^2 they can be represented in the right angle triangle. For example, take f_r^2

$$f_r^2 = <\omega>^2 + B^2. \tag{3.35}$$

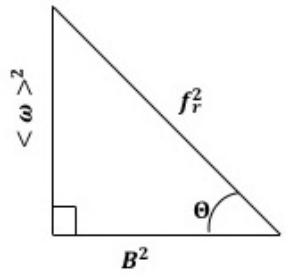


Figure 3-1: Frequency triangle of mean square.

Hence, new feature can be proposed by taking cue from [60]

$$\Theta_{f_r}(t) = \arctan\left[\frac{<\omega>}{B}\right]. \tag{3.36}$$

This feature is defined similar to instantaneous phase. We can also take derivative of this and dividing it by 2π and integrate with density $\int \Theta'_{f_r}(t)a^2(t)$ to get average quantity as it frequency triangle as shown in Figure 3-1. This feature is not used in the work and will be tried upon in near future. From f_r, we can not deduce which component ($B^2, <\omega>^2$) has contributed to it. But from $\Theta_{f_r}(t)$, we can infer which component has contributed to f_r by implying from the $\Theta_{f_r}(t)$.

Chapter 4

Root Mean Instantaneous Frequency Square

4.1 Introduction

After getting good results with RMS frequency, we extend our work with dominant frequency and root mean instantaneous frequency square. In this chapter, we discuss the new features along with RMS frequency. We followed the same methodology, i.e. applying EMD to the Bonn dataset then after Hilbert transform to bring the real valued IMF to complex domain where we can calculate instantaneous frequency $\varphi'(t)$ and amplitude $a(t)$. We calculate the center frequency $< \omega >$ and bandwidth square B^2 or σ_ω^2 using Hermitian time frequency operator as described in previous chapter. In this chapter, we introduce equivalent of RMS frequency that is called dominant frequency. These frequencies are equal but are different in their definition. We define new feature, i.e root mean instantaneous frequency based on the relationship between RMS frequency and dominant frequency.

4.2 Dominant Frequency

Using Parseval's theorem, $\int |S(\omega)|^2 d\omega = \int |s(t)|^2 dt$. Total energy over all the frequencies i.e. $\int |S(\omega)|^2 d\omega$ or directly calculated from time waveform $\int |s(t)|^2 dt$ is

normalized to one. Root mean square frequency is given by [60] and from equation (3.2b)

$$f_r = \sqrt{< \omega >_S^2 + B^2}, \tag{4.1a}$$

$$f_r = \sqrt{< \omega >_T^2 + B_{AM}^2 + B_{FM}^2} \tag{4.1b}$$

Taking cue form [61], we define parameters ratio or contributio ratio as

$$\Gamma_{RMS} = \frac{< \omega >_T^2}{B^2}. \tag{4.2}$$

4.3 Root Mean Instantaneous Frequency Square

Now extending the concepts already discuss in previous chapter, we define dominant frequency. By refering [70], when we expand B_{FM}^2 in equation (3.9f), we get

$$\sigma_\omega^2 = \int a'^2(t)dt + \int (\varphi'(t) - < \omega >)^2 a^2(t)dt, \tag{4.3a}$$

$$\sigma_\omega^2 = \int a'^2(t)dt + < \omega^2 >_T - < \omega >_T^2, \tag{4.3b}$$

$$\sigma_\omega^2 = \int a'^2(t)dt + < \omega^2 >_T - < \omega >_S^2, \tag{4.3c}$$

$$\sigma_\omega^2 + < \omega >_S^2 = \int a'^2(t)dt + < \omega^2 >_T, \tag{4.3d}$$

$$< \omega^2 >_S = \int a'^2(t)dt + < \omega^2 >_T . \tag{4.3e}$$

As $< \omega >_S = < \omega >_T$ and $< \varphi'^2(t) >_T = < \omega^2 >_T$. Now, to confirm the above equation (4.3e), average square frequency or mean square frequency [62] can be expressed by applying time frequency operator, we called it dominant frequency as Barnes et al. [60]

$$< \omega^2 >_S = \int \omega^2 |S(\omega)|^2 d\omega, \tag{4.4a}$$

$$= \int s^*(t) \left(\frac{1}{j} \frac{d}{dt} \right)^2 s(t)dt, \tag{4.4b}$$

$$= - \int s^*(t) \frac{d^2}{dt^2} s(t)dt, \tag{4.4c}$$

$$= \int \left| \frac{d}{dt} s(t) \right|^2 dt, \tag{4.4d}$$

$$= \int \Big(\frac{a'(t)}{a(t)} \Big)^2 a^2(t) dt + \int \varphi'^2(t) a^2(t) dt, \tag{4.4e}$$

$$= \int a'^2(t) dt + < \varphi'^2(t) >_T, \tag{4.4f}$$

$$< \omega^2 >_S = B_{AM}^2 + < \omega^2 >_T, \tag{4.4g}$$

$$f_d = \sqrt{B_{AM}^2 + < \omega^2 >_T} \tag{4.4h}$$

We call f_d as dominant frequency (DF). Equation (4.4g) shows spectral averaged frequency square $< \omega^2 >_S$ is always greater time averaged instantaneous frequency square $< \omega^2 >_T$ (MIFS) by average of square instantaneous bandwidth $\sigma_\omega^2(t)$. Features are presented in Figure 4-2.

Fractional contribution to dominant frequency and their parameters, B_{AM}^2 and $< \varphi'^2(t) >_T$ represented as $< \omega^2 >_T$ ratio can be given by

$$\delta_{DF_{AM}} = \frac{B_{AM}^2}{f_d}, \tag{4.5}$$

$$\delta_{DF_{<\omega^2>_T}} = \frac{< \omega^2 >_T}{f_d}, \tag{4.6}$$

$$\Delta_{DF} = \frac{< \omega^2 >_T}{B_{AM}^2}. \tag{4.7}$$

By substituting resultant equations from (5.4), (3.9f), (4.4f) into (3.2a) as $f_d^2 = f_r^2$, we get

$$\int a'^2(t) dt + \int \varphi'^2(t) a^2(t) dt = \int a'^2(t) dt + \int (\varphi'(t) - < \omega >)^2 a^2(t) dt + \Big(\int \varphi'(t) a^2(t) dt \Big)^2, \tag{4.8}$$

By cancelling out $B_{AM}^2 = \int a'^2(t) dt$, we get

$$\int \varphi'^2(t) a^2(t) dt = \int (\varphi'(t) - < \omega >)^2 a^2(t) dt + \Big(\int \varphi'(t) a^2(t) dt \Big)^2, \tag{4.9a}$$

$$< \varphi'^2(t) >_T = \sigma_T^2 + < \omega >_T^2, \tag{4.9b}$$

$$< \omega^2 >_T = B_{FM}^2 + < \omega >_T^2, \tag{4.9c}$$

$$f_R = \sqrt{B_{FM}^2 + < \omega >_T^2}. \tag{4.9d}$$

f_R is root mean instantaneous frequency square (RMIFS) expressed as root over sum of instantaneous frequency spread around center frequency and square of center frequency. Equation (4.9a) can also be derived by expanding B_{FM}^2.

Fractional contribution to RMIFS and their parameters, B_{FM}^2 and $< \omega >_T^2$ ratio can be given by

$$\varepsilon_{RMIFS_{FM}} = \frac{B_{FM}^2}{f_R}, \tag{4.10}$$

$$\varepsilon_{RMIFS_{<\omega>_T^2}} = \frac{< \omega >_T^2}{f_R} \tag{4.11}$$

$$E_{RMIFS} = \frac{< \omega >_T^2}{B_{FM}^2}. \tag{4.12}$$

Substituting the values from equation (4.9c) to $\delta_{DF_{<\omega^2>_T}}$ and Δ_{DF}

$$\delta_{DF_{<\omega^2>_T}} = \frac{< \omega^2 >_T}{f_d} = \frac{B_{FM}^2 + < \omega >_T^2}{f_d}, \tag{4.13}$$

$$\Delta_{DF} = \frac{B_{FM}^2 + < \omega >_T^2}{B_{AM}}. \tag{4.14}$$

Therefore $\delta_{DF_{<\omega^2>_T}}$ can be factored into

$$\delta_{DF_{FM}} = \frac{B_{FM}^2}{f_d} = \frac{B_{FM}^2}{f_r} = \gamma_{RMS_{FM}}, \tag{4.15}$$

$$\delta_{DF_{<\omega>_T^2}} = \frac{< \omega >_T^2}{f_d} = \frac{< \omega >_T^2}{f_r} = \gamma_{RMS_{<\omega>_T^2}}. \tag{4.16}$$

And Δ_{DF} can also be expressed in terms of

$$\Delta_{DF^{\mp}} = \frac{B_{FM}^2}{B_{AM}^2}, \tag{4.17}$$

$$\Delta_{DF^{\pm}} = \frac{< \omega >_T^2}{B_{AM}^2}. \tag{4.18}$$

Hence, $\Delta_{DF^{\mp}}$ can be called as bandwidth contribution ratio $(\tau_B)^{-1}$ and $\Delta_{DF^{\pm}} = (\Gamma_{RMS_{AM}})^{-1}$. Similar to Θ_{f_r} feature, other features can also be derived as in Figure 4-1.

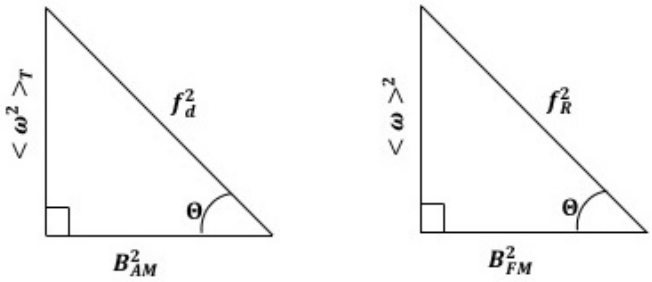


Figure 4-1: Frequency triangle of mean square (dominant frequency) and MIFS.

$$\Theta_{f_d}(t) = \arctan\left[\frac{\sqrt{< \omega^2 >_T}}{B_{AM}}\right], \tag{4.19}$$

$$\Theta_{f_R}(t) = \arctan\left[\frac{< \omega >}{B_{FM}}\right]. \tag{4.20}$$

4.4 SVM

This time we have used RBF and polynomial kernel only. The kernel functions used by us in this work are as follows :

Radial Basis function (RBF) kernel is represented by

$$G(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(\frac{||\boldsymbol{x_i} - \boldsymbol{x_j}||^2}{2\sigma^2}\right), \tag{4.21}$$

where σ is a positive number.

Polynomial Kernel can be represented by

$$G(\boldsymbol{x}_i, \boldsymbol{x}_j) = (1 + \boldsymbol{x}_i^T \boldsymbol{x}_j)^p, \tag{4.22}$$

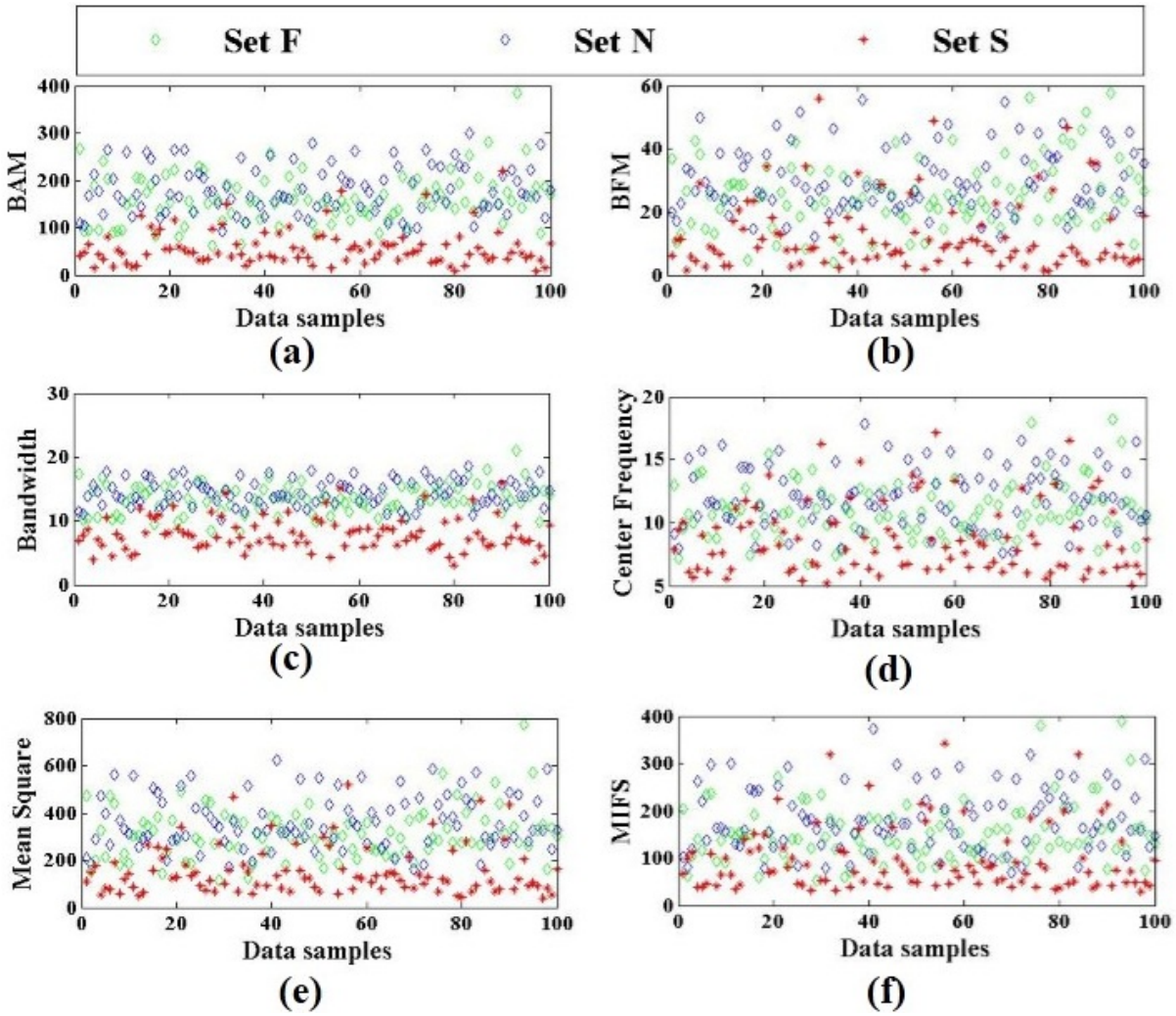


Figure 4-2: Features from IMF2 (a) Bandwidth amplitude modulation B_{AM}^2, (b) Bandwidth frequency modulation B_{FM}^2, (c) Root mean bandwidth B from equation (3.9h), (d) Center frequency $< \omega >_T$, (e) Mean square frequency f_r^2 from equation (3.2a), (f) Mean instantaneous frequency square $< \varphi'^2(t) >_T$.

where p is the degree of polynomial.

4.5 Material and Simulation

4.5.1 Dataset

We have applied our method on EEG dataset [65] commonly known as Bonn dataset. In our work subset F, N and S are used by us as they are all results of intracranial experiments and have not used subset Z, O as they were extracranially recorded. We have created three classes for classification. One with sets F, S, second with N, S and third with F, N against S.

Each sets contain 100 data from 100 signal segments for each feature. These data are normalized using standard deviation and mean. We have prepared the training and test set in 70-30 ratio. One from interictal (F) and ictal (S), second from interictal (N) and ictal (S). We took 70 data samples randomly without replacement from each interictal set and ictal set. For test set, remaining 30 data from interictal and ictal sets are taken. This way we have total 140 data from ictal and interictal set for training and 60 data samples from ictal and interictal set for testing. SVM was applied on these sets. We have tried to optimise the parameters of RBF kernel using grid serch and found default parameters of $\sigma = 1.0$ and $Penalty = 1.0$ are best with these features. We repeated this process hundred times means taking 100 trials following Bajaj et al. [36] who has taken 10 trials. Results and figures with default kernel parameters are presented in Table. 4.3-4.8 and in Figure 4-7-4-6.

For classification of set F, N against set S, we have randomly picked 35 data samples from each F and S set for training. For creating test set, we picked randomly 15 data samples from remaining 65. From set S, we picked randomly 70 data samples for training and 30 remaining been taken for creating test set. This way taking 70 data samples each from interictal set F, N and ictal set S equally for training set. Picking equal number of data samples from interictal and from ictal set helps in avoiding over fitting.

4.5.2 Simulation

We applied Hilbert transform to the real valued IMFs which take it to time frequency energy complex domain. It enable us to calculate the instantaneous frequency [54]. When B_{AM}^2 and B_{FM}^2 in B^2 computed, both these component were in higher range in interictal set compare to ictal sets as represented in the Figure 4-2. Time frequency operator was used to compute these and rest of the features which when applied to spectral feature, gives us time conditional components. The frequency operator $\left(\frac{1}{j}\frac{d}{dt}\right)$ takes the feature from spectral domain to time frequency domain, giving the time information, when it happened. For example, when frequency operator was applied on spectral bandwidth σ_ω^2, i.e. spread of frequencies in (3.9a), it gives an output containing average instantaneous bandwidth square and spread of instantaneous frequencies around center frequency as in (3.9e). These two quantities shows local spread at instant of time, i.e. $\sigma_\omega(t)$ and deviation or spread of instantaneous frequencies around the global mean or center frequency. From B_{AM}^2, we can infer three quantities, one as average square local spread, second total amplitude modulation as in equation 3.9h [61] and third, it can be taken as local instantaneous energy [71]. Equation (3.9g) represents spectral bandwidth square is always greater than the global spread of instantaneous frequencies by average square local spread.

Similarly, when we apply frequency operator on second moment $< \omega^2 >_S$, it can be expressed as summation of instantaneous bandwidth square $\sigma_\omega^2(t)$ and instantaneous frequency square $< \varphi'^2(t) >$. Integrating with density function over all time gives their average quantity or amplitude weighted quantity. By taking root over, dominant frequency f_d is interpreted as square root of sum of square of instantaneous bandwidth and square of instantaneous frequency $< \varphi'^2(t) >$ or $< \omega^2 >_T$. $< \omega^2 >_S$ is always greater than $< \omega^2 >_T$ by average of square of instantaneous bandwidth $\sigma_\omega^2(t)$. Therefore, $< \omega^2 >_S \neq < \omega^2 >_T$ i.e. spectral mean square is not equal to time averaged mean square.

Instantaneous bandwidth or local spread $\sigma_\omega(t)$ gives information about the range of frequencies changing with respect to instant of the time. For example, commonly

average height and the standard deviation of population of area is known. But what if we want to know standard deviation of height at birth, at age of 3. Local spread gives this information [61].

Now, instantaneous bandwidth $\sigma_\omega(t)$ has a drawback that for given two different signals, their instantaneous bandwidth can be same [72]. For example, consider two signals $s_1 = a_1(t)e^{j\varphi(t)}$ and $s_2 = a_2(t)e^{j\varphi(t)}$ where $a_2 = \frac{1}{a_1(t)}$ and $\varphi(t)$ representing phase function. Then, instantaneous bandwidth square of s_2, as in [72] can be given by

$$\left(\frac{a_2'(t)}{a_2(t)}\right)^2 = \left(\frac{\frac{d}{dt}\frac{1}{a_1(t)}}{\frac{1}{a_1(t)}}\right)^2 = \left(\frac{\frac{-a_1(t)}{a_1^2(t)}}{\frac{1}{a_1(t)}}\right)^2 = \left(\frac{a_1'(t)}{a_1(t)}\right)^2 \tag{4.23}$$

is identical to instantaneous bandwidth square of s_1. Now, B_{FM}^2 in B^2 mask B_{AM}^2 and its drawback. But B^2 of more than two signals can also be same. Although signals can have different AM and FM contribution yet they may end up with same bandwidth [62].

To overcome these drawbacks, one way is to use instantaneous kurtosis [72],[69], [73] . But we could not achieve significant difference between ictal and interictal with it. RMS frequency can be used as feature for classification purpose with an advantage of overcoming aforementioned drawback. Adding square of center frequency to square bandwidth creates unique feature. Similarly, we can use dominant frequency for classification where average instantaneous frequency square is added to instantaneous bandwidth square. It can overcome the drawback of instantaneous bandwidth too. Though, in Bonn dataset, for ictal and interictal sets we have not encountered exactly same values of B_{AM}^2, B^2 yet this features remain prone to mentioned drawback. However, there are examples present with non seizure and seizure signal. We present the solution in the Figure 4-3-4-4

From Figure 4-2, we observe that along with B_{AM}^2, B_{FM}^2, root mean bandwidth B, center frequency $< \omega >_T$, mean square frequency f_r^2, f_d^2 and mean instantaneous frequency f_R^2 or $< \varphi'^2(t) >_T$ are in higher range in interictal than ictal signals. Whereas in Figure 4-5, we can observe that the parameters ratio of f_r, f_d and f_R shows higher values for ictal than interictal. These contributing parameters ratio shows how dom-

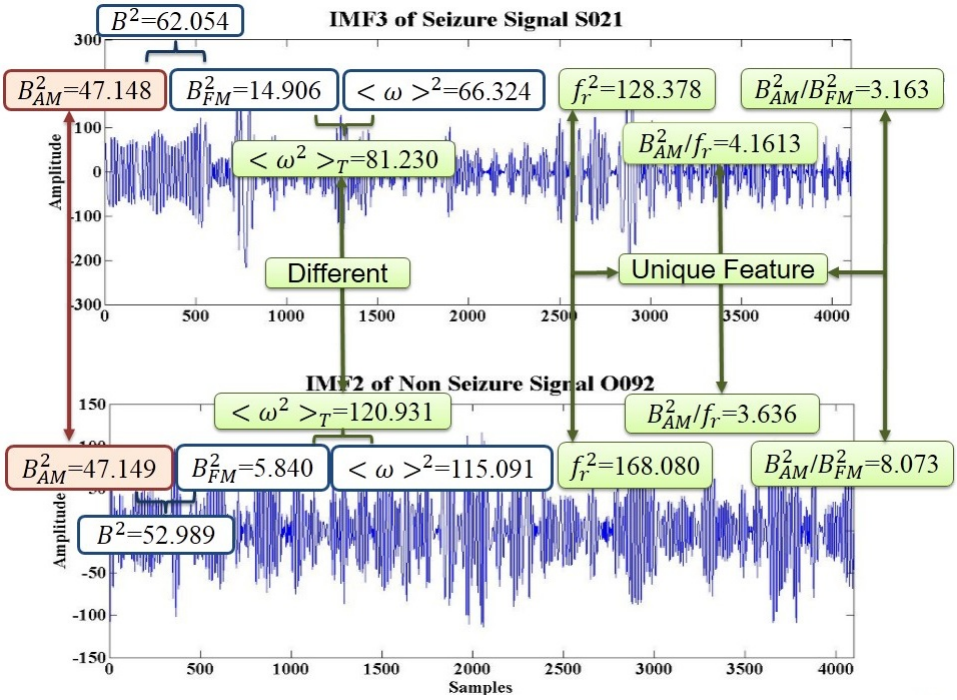


Figure 4-3: Similar B^2_{AM}.

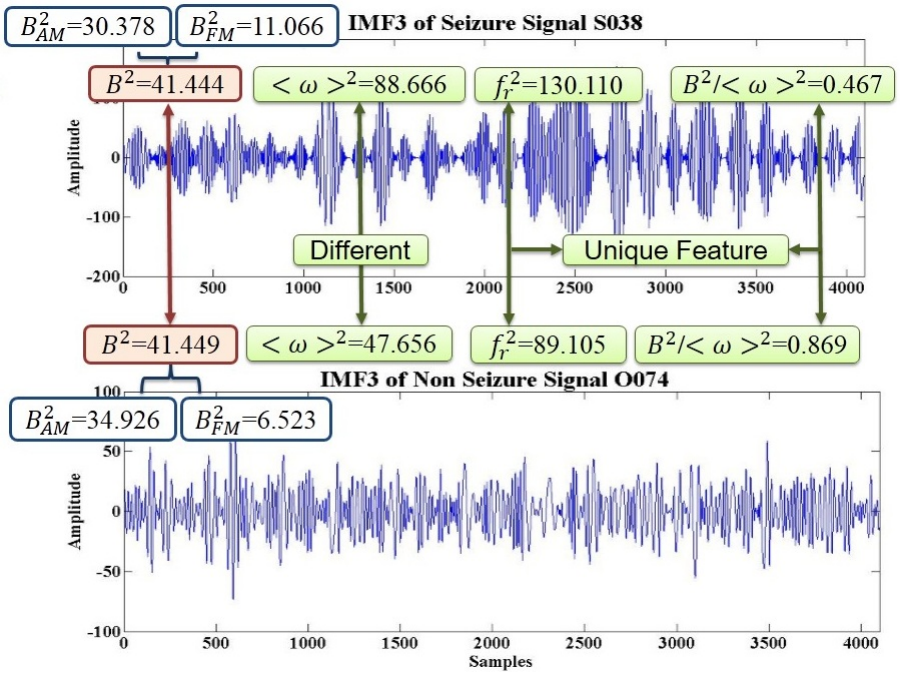


Figure 4-4: Similar B^2.

inant is the frequency component over the bandwidth term in the feature at instant of time or for an event as in ictal in these cases. For example, if we take Γ_{RMS} i.e. $< \omega >_T^2 / B^2$ which is at higher range in ictal than interictal means $< \omega >^2$ is much greater than B^2 in ictal, whereas there is not much difference between them in interictal. From these observation, we can infer that there is large bandwidth in interictal than ictal due to large modulations in interictal than ictal. After analysing the features, their contribution ratios and their parameters ratios we have chosen the f_r, f_d, f_R features and their parameters ratios $\frac{<\omega>_T^2}{B^2}$, $\frac{<\omega^2>_T}{B_{AM}^2}$, $\frac{<\omega>_T^2}{B_{FM}^2}$ for classification purpose in this work as these features shows opposite behaviour. Though the feature f_r and f_d are equal yet they are different by the contributing parameters and so their parameters ratios $\frac{<\omega>_T^2}{B^2}$, $\frac{<\omega^2>_T}{B_{AM}^2}$. Using the relationship $f_r^2 = f_d^2$, we have defined $< \varphi'(t)^2 >_T = \sigma_T^2 + < \omega >_T^2$. By taking root over, we called it root mean instantaneous frequency square (RMIFS), f_R. It follows the definition of RMS as the first term is time averaged standard deviation square σ_T^2 and second term is squared center frequency $< \omega >_T^2$. It does not contain instantaneous bandwidth which makes it free from its drawback. One more advantage is f_R is defined in terms of instantaneous frequency only and totally dependent on phase function $\varphi(t)$ and not on amplitude. Feature f_R values are at higher range in interictal and lower in ictal similar to other features $f_r,\ f_d$.

We have presented the range of RMS frequency and dominant frequency together in Table 4.1 as RMS and dominant frequency have same values. Features in IMF1 are in gamma brain wave range on an average whereas they are in alpha, beta range in IMF2. RMIFS frequency is spread across beta and lower gamma range in IMF1. Overall, on an average it is in beta brain wave range in IMF1 whereas its in beta range in IMF2 for interictal. For ictal it is in beta range on an average in IMF1 and in IMF2 it is spread across theta, alpha and beta wave.

Table 4.1: RMS frequency f_r, Dominant frequency (DF) f_d and RMIFS f_R features range.

Feature	Set	IMF1 min-max avg [Hz]	IMF2 min-max avg [Hz]
RMS (f_r) and DF (f_d)	F	20.50-74.70 46.42	10.79-27.82 17.37
	N	27.99-78.58 54.17	12.78-25.01 18.79
	S	13.46-54.63 21.46	6.182-22.81 11.83
RMIFS (f_R)	F	11.74-53.64 29.88	7.315-19.74 12.04
	N	17.83-55.68 35.42	8.345-19.29 13.22
	S	7.816-36.35 15.64	5.438-18.52 9.293

After taking the $p-values$ from Kruskal-Wallis test between set F, N and S into consideration, we have decided onto features from first two IMFs to feed to SVM for classification. Table 4.2 shows the $p-values$ of first four IMFs. Figure 4-5 shows the features used in this classification work.

The statistical parameters for performance evaluation are used as utilized in previous works i.e. sensitivity (SEN), specificity (SPE) and Accuracy (Acc) [20] [36]. where TP stands for true positive event i.e. detecting ictal correctly. FN stands for false negative. FN signifies the failure to identify ictal which it has labelled as interictal. FP shows false positive. TN represents true negative event.

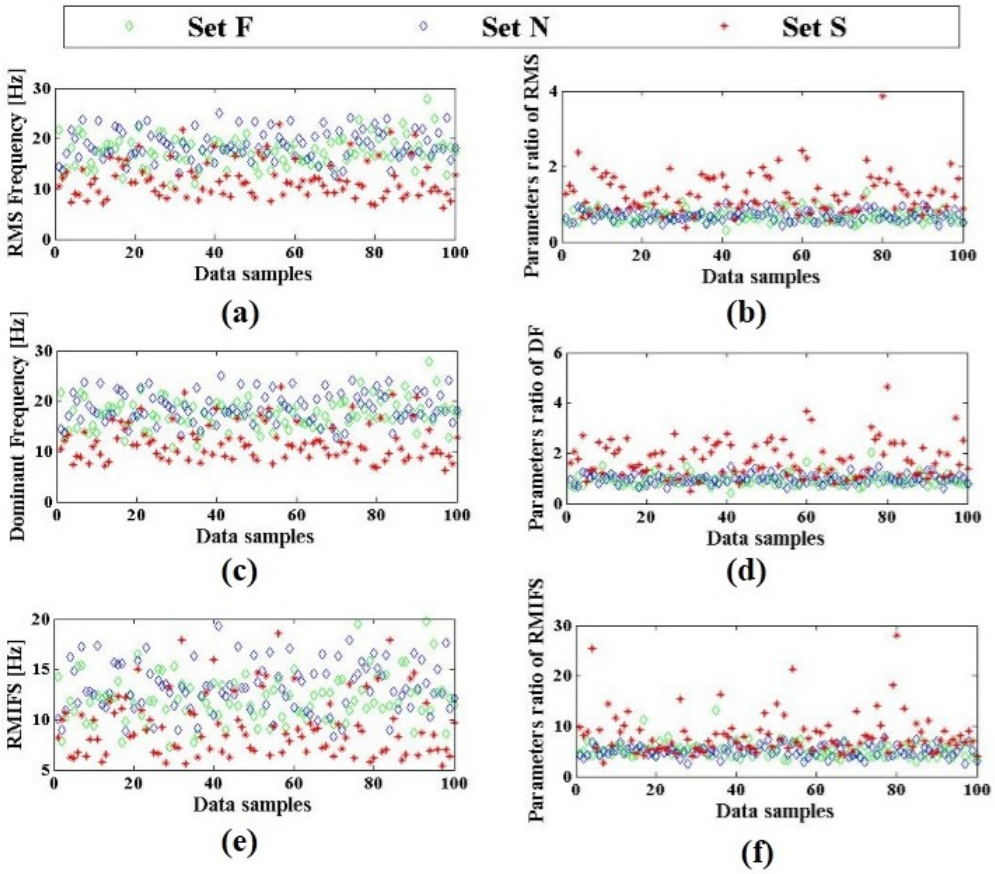


Figure 4-5: RMS frequency, dominant frequency, RMIFS and their parameters ratio from IMF2.

Table 4.2: Kruskal-Wallis test and $p-values$ of features from Bonn dataset subsets F, N and S.

Feature	IMF1	IMF2	IMF3	IMF4
f_r	1.92e-40	1.48e-29	8.94e-16	2.87e-08
Γ_{RMS}	7.40e-36	2.51e-31	2.89e-18	8.74e-09
f_d	1.92e-40	1.92e-40	8.94e-16	2.17e-06
Δ_{DF}	7.26e-28	5.76e-29	2.80e-17	2.86e-07
f_R	5.04e-37	1.92e-40	7.40e-36	2.17e-06
E_{RMIFS}	8.51e-38	1.10e-19	5.63e-13	6.60e-12

4.6 Results and Discussions

Classification of interictal set F from ictal set S (refer Table 4.3) using RMS frequency f_r and its parameters ratio $\frac{<\omega>_T^2}{B^2}$ shows consistent average sensitivity of $99.50\%, 100\%$ and average accuracy of $98.66\%, 98.71\%$ for RBF kernel in IMF1-2. Polynomial kernel has shown consistent 100% average sensitivity in both IMF1-2. Classification of set N against set S shows 100% of average sensitivity in IMF1-2 for both RBF kernel and Polynomial kernel. The highest accuracy 99.91% was observed in IMF2 using RBF kernel. Figure 4-6-4-7 shows 100% classification using default kernel parameters.
We have observed similar results using f_d and it parameters ratio $\frac{<\omega^2>_T}{B_{AM}^2}$. Though slightly less compare to RMS frequency and its parameters ratio yet consistent for both RBF kernel and polynomial kernel refer Table 4.5-4.6. Classification of set F against set S shows average sensitivity of $99.80\%, 100\%$ for RBF kernel and $96.56\%, 98.33\%$ using polynomial kernel for IMF1-2. Average accuracy in IMF1 for RBF and Polynomial kernel were $96.33\%, 96.43\%$ whereas in IMF2 it is observed as $98.51\%, 97.15\%$ for both the kernel. Classification of set S from N shows better results of average sensitivity of 100% in IMF1-2 for RBF kernel and $99.43\%, 99.56\%$ for polynomial kernel. Highest accuracy observed in IMF2 of $99.75\%, 99.38\%$ for both kernel.
Classification with RMIFS f_R and its parameters ratio $\frac{<\omega>_T^2}{B_{FM}^2}$ shows good results in IMF1 than IMF2 refer Table 4.7-4.8 as against root mean square frequency f_r and dominant frequency f_d which shows better result in IMF2. Average sensitivity in IMF1-2 was observed as $99.36\%, 98.43\%$ and average accuracy was $98.30\%, 95.10\%$ in IMF1-2 for classification of set F and set S for RBF kernel. Similar results were observed for Polynomial kernel as shown in Table 4.7. Classification of set N vs S has average accuracy of $98.71\%, 98.18\%$ in IMF1-2 for RBF Kernel. Using Polynomial kernel, average accuracy observed was 98.20% in IMF1 and 97.31% in IMF2.
Next, we have done classification of interictal set F, N together against ictal set S. Refer Table 4.9, using RMS frequency f_r and its parameters ratio $\frac{<\omega>_T^2}{B^2}$, we could attain 98.05% of average accuracy with 100% of average sensitivity by employing RBF kernel in IMF2.

Table 4.3: Result of classification with RMS frequency f_r and its ratio of parameters Γ_{RMS} on interictal (F) and ictal (S) of Bonn dataset for 100 trials.

Kernel Function	Statistical parameters	IMF1 min-max avg [%]	IMF2 min-max avg [%]
RBF ($\sigma = 1.0$, $Penalty = 1.0$)	SEN	93.33-100.0 99.50	100.0-100.0 100.0
	SPE	80.00-100.0 97.83	86.66-100.0 97.43
	Acc	90.00-100.0 98.66	93.33-100.0 98.71
Polynomial ($p = 3$)	SEN	100.0-100.0 100.0	100.0-100.0 100.0
	SPE	70.00-96.66 84.50	86.66-100.0 93.46
	Acc	85.00-98.33 92.25	93.33-100.0 96.73

Table 4.4: Result of classification with RMS frequency f_r and its ratio of parameters Γ_{RMS} on interictal (N) and ictal (S) of Bonn dataset for 100 trials.

Kernel Function	Statistical parameters	IMF1 min-max avg [%]	IMF2 min-max avg [%]
RBF ($\sigma = 1.0$, $Penalty = 1.0$)	SEN	100.0-100.0 100.0	100.0-100.0 100.0
	SPE	93.33-100.0 99.06	96.66 100.0 99.83
	Acc	96.66-100.0 99.53	98.33-100.0 99.91
Polynomial ($p = 3$)	SEN	100.0-100.0 100.0	100.0-100.0 100.0
	SPE	63.33-90.00 77.53	86.66-100.0 93.80
	Acc	81.66-95.00 88.76	93.33-100.0 96.90

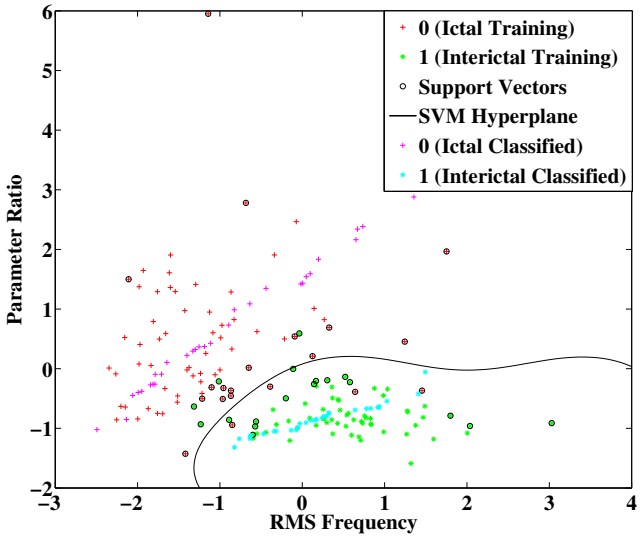


Figure 4-6: Classification of interictal set (F) vs ictal set (S) from IMF2 using RBF kernel.

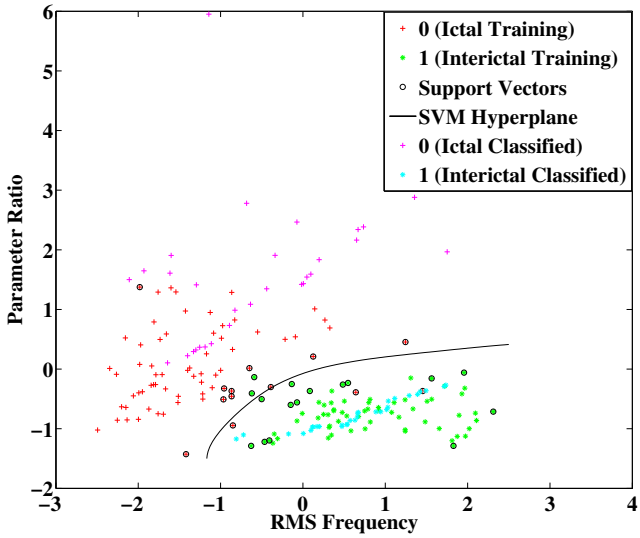


Figure 4-7: Classification of interictal set (F) vs ictal set (S) from IMF2 using RBF kernel.

Table 4.5: Result of classification with dominant frequency f_d and its ratio of parameters Δ_{DF} on interictal (F) and ictal (S) of Bonn dataset for 100 trials.

Kernel Function (Default Parameters)	Statistical parameters	IMF1 min-max avg [%]	IMF2 min-max avg [%]
RBF ($\sigma = 1.0$, $Penalty = 1.0$)	SEN	93.33-100.0 99.80	100.0-100.0 100.0
	SPE	73.33-100.0 92.86	86.66-100.0 97.03
	Acc	86.66-100.0 96.33	93.33-100.0 98.51
Polynomial ($p = 3$)	SEN	76.66-100.0 96.56	86.66-100.0 98.33
	SPE	76.66-100.0 96.30	76.66-100.0 95.46
	Acc	88.33-100.0 96.43	85.00-100.0 97.15

Table 4.6: Result of classification with dominant frequency f_d and its ratio of parameters Δ_{DF} on interictal (N) and ictal (S) of Bonn dataset for 100 trials.

Kernel Function (Default Parameters)	Statistical parameters	IMF1 min-max avg [%]	IMF2 min-max avg [%]
RBF ($\sigma = 1.0$, $Penalty = 1.0$)	SEN	100.0-100.0 100.0	100.0-100.0 100.0
	SPE	93.33-100.0 98.30	93.33-100.0 99.50
	Acc	96.66-100.0 99.15	96.66-100.0 99.75
Polynomial ($p = 3$)	SEN	93.33-100.0 99.43	90.00-100.0 99.56
	SPE	90.00-100.0 98.83	90.00-100.0 99.20
	Acc	95.00-100.0 99.13	95.00-100.0 99.38

Table 4.7: Result of classification with RMIFS frequency f_R and its ratio of parameters E_{RMIFS} on interictal (F) and ictal (S) of Bonn dataset for 100 trials.

Kernel Function (Default Parameters)	Statistical parameters	IMF1 min-max avg [%]	IMF2 min-max avg [%]
RBF ($\sigma = 1.0$, $Penalty = 1.0$)	SEN	86.66-100.0 99.36	66.66-100.0 98.43
	SPE	80.00-100.0 97.23	70.00-100.0 91.76
	Acc	83.33-100.0 98.30	80.00-100.0 95.10
Polynomial ($p = 3$)	SEN	83.33-100.0 97.73	40.00-100.0 89.96
	SPE	83.33-100.0 96.86	66.66-100.0 92.76
	Acc	86.66-100.0 97.30	66.66-100.0 91.33

Table 4.8: Result of classification with RMIFS frequency f_R and its ratio of parameters E_{RMIFS} on interictal (N) and ictal (S) of Bonn dataset for 100 trials.

Kernel Function (Default Parameters)	Statistical parameters	IMF1 min-max avg [%]	IMF2 min-max avg [%]
RBF ($\sigma = 1.0$, $Penalty = 1.0$)	SEN	96.66-100.0 99.93	80.00-100.0 99.60
	SPE	90.00-100.0 97.50	80.00-100.0 96.76
	Acc	95.00-100.0 98.71	90.00-100.0 98.18
Polynomial ($p = 3$)	SEN	86.66-100.0 98.33	70.00-100.0 98.20
	SPE	90.00-100.0 98.06	73.33-100.0 96.43
	Acc	90.00-100.0 98.20	80.00-100.0 97.31

Table 4.9: Result of classification on interictal (F, N) and ictal (S) of Bonn dataset using default parameters of RBF kernel ($\sigma = 1.0$, $Penalty = 1.0$) and polynomial kernel ($p = 3$) for 100 trials.

Feature	Kernel Function (Default Parameters)	Statistical parameters	IMF1 min-max avg [%]	IMF2 min-max avg [%]
RMS frequency f_r and its parameters ratio	RBF	SEN	96.66-100.0 99.90	100.0-100.0 100.0
		SPE	76.67-100.0 92.60	80.00-100.0 96.10
		Acc	88.33-100.0 96.25	90.00-100.0 98.05
	Polynomial	SEN	90.00-100.0 98.66	90.00-100.0 99.50
		SPE	86.66-100.0 94.86	83.33-100.0 97.10
		Acc	91.66-100.0 96.76	91.66-100.0 98.30
Dominant frequency f_d and its parameters ratio	RBF	SEN	93.33-100.0 99.93	100.0-100.0 100.00
		SPE	80.00-100.0 91.93	83.33-100.0 95.80
		Acc	90.00-100.0 95.93	91.66-100.0 97.90
	Polynomial	SEN	86.66-100.0 98.36	93.33-100.0 99.50
		SPE	76.66-100.0 94.23	76.66-100.0 95.76
		Acc	83.33-100.0 96.30	88.33-100.0 97.63
RMIFS frequency f_R and its parameters ratio	RBF	SEN	93.33-100.0 99.67	73.33-100.0 99.20
		SPE	76.66-100.0 91.03	73.33-100.0 89.73
		Acc	88.33-100.0 95.35	83.33-100.0 94.75
	Polynomial	SEN	90.00-100.0 98.36	56.66-100.0 94.46
		SPE	76.66-100.0 90.93	66.66-100.0 90.83
		Acc	86.66-100.0 94.65	75.00-100.0 92.65

Using RMIFS frequency f_R and its parameters ratio, the results are less compare to using RMS and dominant frequency. Highest average accuracy was 95.35% in IMF1 with RBF kernel showing average sensitivity of 99.67%. The reason for this results is due to false positive as we can see in Table 4.9 that average specificity on an average is $91.03\%, 89.73\%, 90.93\%$ and 90.83% in IMF1-2 using both kernels. Low specificity is due false positives which was pulling down the accuracy.

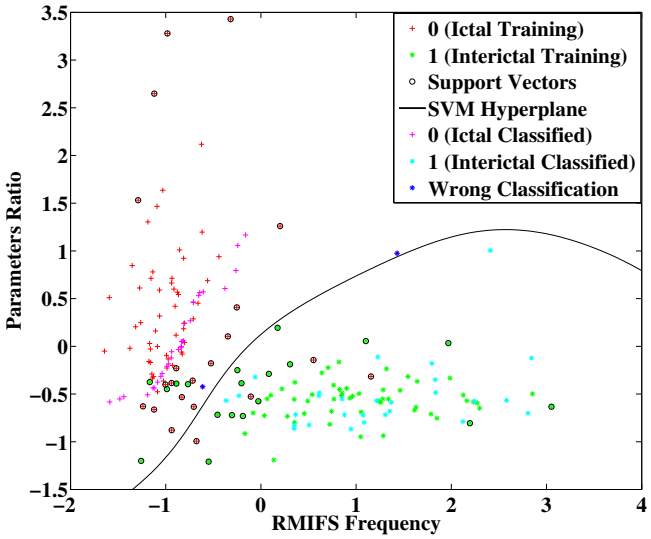


Figure 4-8: Classification of ictal set S vs interictal set F, N from IMF1 using RBF kernel ($\sigma = 1.0$, $Penalty = 1.0$) with wrong classifications.

In the Figure 4-8, we can see the wrong classifications using RMIFS frequency and its parameters ratio $\frac{<\omega>_T^2}{B_{FM}^2}$. We have tried to improve the results with kernel parameters optimization (refer Table 4.10) with RMIFS frequency but could achieve only little improvement. So, we have devised an algorithm based on adaptive thresholding to recover the false positives and consequently the average accuracy. After analysing all the features, fractional contributions, we have decided to use $\delta_{DF_{AM}} = \frac{B_{AM}^2}{f_d}$ from IMF2 as threshold as it is showing greater significant difference compared to others. Using $\delta_{DF_{AM}}$ has an added advantage that it do not contain the drawback of B_{AM}^2 as it is normalized by f_d. We have tried this method with RMS frequency and dominant frequency also. It increases the specificity at the cost of decreased sensitivity which is not desirable. We found this method is suitable where there is good

difference between sensitivity and specificity in higher beta, lower gamma frequency range as in case of IMF1 of RMIFS frequency for the classification of set F, N versus set S. Larger the difference between sensitivity and specificity, better the improvement in identifying the false positive as in our case. The sensitivity slightly decreased with this thresholding method because there are few instance where data samples $\frac{B_{AM}^2}{f_d}$ of ictal are in interictal range (refer Figure 4-9). This samples get relabelled as interictal. By doing so, increasing the the false negative and decreasing the sensitivity. Overall, in IMF1 for RBF (refer Table 4.10), average specificity increased by 5.9% to 96.93% from 91.03% consequently increasing the average accuracy by 2.88% from 95.35% to 98.23%. Average sensitivity decreased by only 0.14%. In polynomial kernel, the average specificity increased by 5.07% from 90.93% to 96.00% which increased the average accuracy by 2.4% from 94.65% to 97.05%. Average sensitivity decreased by 0.26% on an average. In IMF2 there is little improvement using RBF and polynomial polynomial kernel as the difference between sensitivity and specificity is small compared to IMF1 and the frequency range is in theta, alpha and lower beta brain wave.

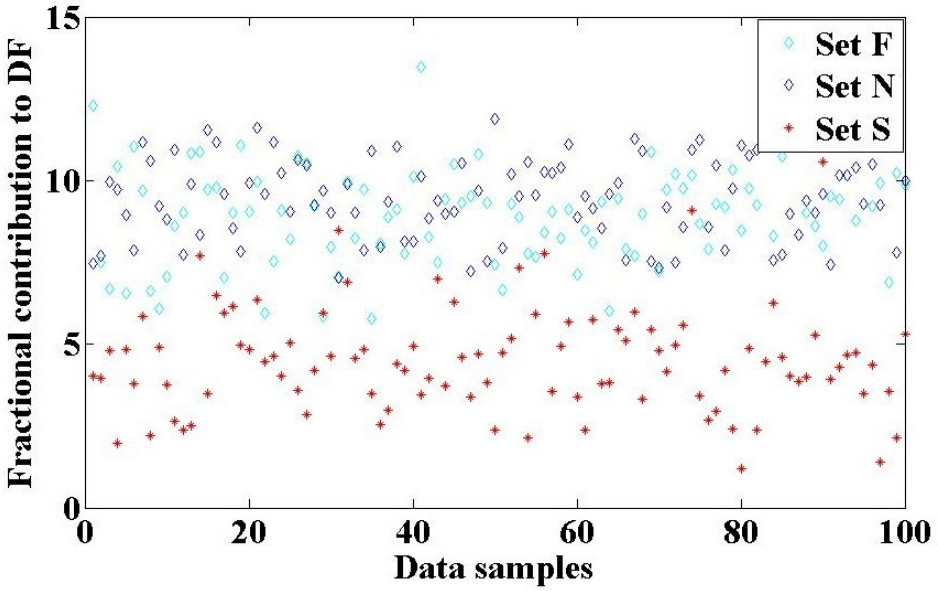


Figure 4-9: Fractional contribution to DF, $\delta_{DF_{AM}} = \frac{B_{AM}^2}{f_d}$ from IMF2 .

The adaptive thresholding algorithm is as follows

Input : Features RMIFS frequency f_R and $\frac{<\omega>_T^2}{B_{FM}^2}$ calculated from first 2 IMF and threshold $\frac{B_{AM}^2}{f_d}$ of IMF2 from set F, N, S.

for *i:=1 to 2* **do**

A=load all features from IMF (i) and threshold $\frac{B_{AM}^2}{f_d}$.

for *trials:=1 to 100* **do**

1. Create train set train for SVM of RMIFS frequency f_R and $\frac{<\omega>_T^2}{B_{FM}^2}$.
2. Create another set threshold-train corresponding to train from $\frac{B_{AM}^2}{f_d}$.
3. Create test set test from remaining data for SVM of RMIFS frequency f_R and $\frac{<\omega>_T^2}{B_{FM}^2}$.
4. Create another set threshold-test corresponding to test from remaining $\frac{B_{AM}^2}{f_d}$.
5. Set threshold as minimum of interictal set F, N of threshold-train in step 2.
6. Create SVM model using train set train with default kernel parameters.
7. Apply test to SVM model.
8. Take all the samples identified as ictal or true positive by SVM model output (OP).
9. Take all the corresponding samples of OP from threshold-test
10. Apply threshold from step 5 to find out their classes.
11. Update OP with new labels from above step 10.
12. Calculate SEN, SPE, Acc of SVM.

end

Take the average of 100 *trials*.

end

Algorithm 1: Adaptive thresholding algorithm.

Table 4.10: Result of classification with RMIFS frequency f_R and its ratio of parameters E_{RMIFS} on interictal (F, N) and ictal (S) of Bonn dataset for 100 trials with optimized kernel parameters and with threshold $\delta_{DF_{AM}} = \frac{B^2_{AM}}{f_d}$.

Kernel Function (Parameters)	Statistical parameters	IMF1 min-max avg [%]	IMF2 min-max avg [%]
RBF ($\sigma = 0.93$, $Penalty = 1.0$)	SEN	93.33-100.0 99.76	80.00-100.0 94.38
	SPE	70.00-100.0 90.40	70.00-100.0 89.33
	Acc	85.00-100.0 95.08	85.00-100.0 94.43
RBF ($\sigma = 1.0$, $Penalty = 1.0$) with threshold	SEN	93.33-100.0 99.53	70.00-100.0 98.86
	SPE	80.00-100.0 96.93	76.66-100.0 92.76
	Acc	90.00-100.0 98.23	85.00-100.0 95.81
Polynomial ($p = 3$) with threshold	SEN	86.66-100.0 98.10	56.66-100.0 95.56
	SPE	80.00-100.0 96.00	73.33-100.0 92.93
	Acc	90.00-100.0 97.05	75.00-100.0 94.25

We have compared our work to previous related works , refer Table 5.3-5.4 and found the performance of RMS frequency, dominant frequency and RMIFS frequency good. Features RMS frequency f_r and dominant frequency f_d from IMF2 are in alpha, beta range of brain wave which we consider more meaningful and having physical relevance than the features from IMF1, exhibiting high gamma brain wave range.

Table 4.11: Comparison with other works on classification of subsets F vs S and N vs S of Bonn dataset.

Author	Decomposition or Preprocessing Method	Features Used	Classifier	Set	Accuracy
Liang et al. [20]	Fast Fourier Transform	16 spectral features	SVM	F vs S	98.74%
Nicolau et al. [25]	NA	Permutation entropy	SVM	F vs S	83.13%
				N vs S	88.83%
Siuly et al. [74]	Clustering	9 temporal features	LS-SVM	F vs S	93.91%
				N vs S	97.69%
Zhu et al. [26]	NA	strength and degree of	K-NN	F vs S	93.00%
		HVG features		N vs S	98.00%
Riaz et al. [34]	EMD	6 temporal and	Decision trees	F vs S	96.00%
		spectral features	SVM	F vs S	93.00%
Samiee et al. [75]	RDSTFT	5 time frequency	MLP	F, S	94.90%
		features		N vs S	98.50%
Hassan et al. [30]	CEEMDAN	6 spectral features	Boosting	F vs S	97.00%
				N vs S	100.0%
			SVM	F vs S	93.00%
				N vs S	99.00%
Proposed work	**EMD**	**RMS frequency and its parameters ratio**	**SVM**	**F vs S**	**98.71%**
				N vs S	**99.91%**
Proposed work	**EMD**	**Dominant frequency and its parameters ratio**	**SVM**	**F vs S**	**98.51%**
				N vs S	**99.75%**
Proposed work	**EMD**	**RMIFS frequency and its parameters ratio**	**SVM**	**F vs S**	**98.30%**
				N vs S	**98.71%**

Note: Not applicable (NA), rational discrete STFT (RDSTFT), complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN)

Table 4.12: Comparison with other works on classification of subsets F, N vs S of Bonn dataset.

Author	Decomposition or Preprocessing Method	Features Used	Classifier	Set	Accuracy
Sharma et al. [37]	EMD	2D, 3D PSR	LS-SVM	F, N vs S	98.67%
Altunay et al. [76]	Linear Predictiver Filter	Energy based feature	Threshold	F, N vs S	94.00%
Joshi et al. [77]	Fractional Linear Prediction	FLP Energy signal energy	SVM	F, N vs S	95.33%
Pachori et al. [78]	EMD	SODP of IMF	ANN	F, N vs S	97.75%
Proposed work	**EMD**	**RMS frequency and its parameters ratio**	**SVM**	**F, N vs S**	**98.30%**
Proposed work	**EMD**	**Dominant frequency and its parameters ratio**	**SVM**	**F, N vs S**	**97.90%**
Proposed work	**EMD**	**RMIFS frequency and its parameters ratio with threshold**	**SVM**	**F, N vs S**	**98.23%**

4.7 Summary

EMD is good in handling and decomposing non stationary EEGs. In this work we have found root mean square frequency f_r or dominant frequency f_d is good aver-

age measure in classifying ictal from interictal EEG with highest average accuracy of 99.91%. These results are obtained from features of IMF2 having range in alpha, beta brain wave. Average sensitivity of 100% was observed using RMS frequency. We have also presented mean instantaneous frequency square in terms of time averaged instantaneous frequency spread and center frequency which is also good in discriminating ictal from interictal with 98.71% of average accuracy. We found selection and combination of features displaying opposite behaviours gives good results. All the features used in this work exhibit best and consistent results with default setting of SVM. This makes them suitable for practical clinical trial as it reduce the computational overload of optimization of SVM kernel parameters. In future we will try to find better feature to address the problem of false positive using our proposed adaptive thresholding method. We will try to utilize rest of the features like fractional contribution of RMS frequency or dominant frequency for classification. Deciding on to which IMF will best classification remains a problem which increases the computational overhead. We will address this issue with some new decomposition method.

Chapter 5

Morphological Component Analysis

In this chapter, We discuss morphological component analysis and its application in epileptic seizure detection. EMD proved good in decomposing the EEG signal but it is computationally heavy because it is not possible to determine feature from which IMF will be good for classification. Although time frequency feature from second IMF in this work are observed to give high classification accuracy yet we will not be sure in future which IMF will be best suited for an application. For example, RMS frequency and dominant frequency has shown good accuracy in IMF2 where as RMIFS frequency has shown best result from IMF1. The reason we found that epileptic classification results are best in beta brave frequency range with these features i.e. feature RMS and dominant frequency from IMF2 and RMIFS frequency from IMF1. But it will remain unknown that certain IMF will lbe in beta range as EMD decompose EEG signal adaptively, depending on the local characteristics of the signal. So applying MCA is suitable as it gives definite number of decomposition and it gives freedom in choosing basis functions according to application.

5.1 Introduction

In epilepsy, the commonly observed behaviour or morphology is spike train, sharp waves. The sudden transient burst of spikes and high frequency oscillation in interictal recording are also used for the localization of the epileptic seizures. Both,

disparity in background activity and EEG paroxysms make the automated analysis complicated. Artifacts in filtered data can give rise to false positive [79] [80] [81]. Recently signal decomposition by focusing morphological components are getting to be highlighted due to the applicability to nonlinear and non-stationary signal properties [82] [83] [84]. Morphological component analysis (MCA) [45] has shown promising results in removing artefacts from EEG. It identify the component of the signal based on sparsity in time frequency domain. MCA decompose the signal and then can accurately reconstruct the signal using redundant transforms (mathematical function) called explicit dictionary. These combination of explicit dictionary form overcomplete dictionary is important for representation of different morphology of EEG signal. Sparse based reconstruction of EEG signal has an advantage of using minimum coefficients which gives it an advantage to be easily transferred it over internet.

This time, we have used MCA with undecimated wavelet transform (UDWT), local discrete cosine transform (LDCT) and Dirac bases composing the dictionary for decomposition. UDWT identify the slow components in the EEG representing mainly eye artefacts, LDCT identify the spectral components and Dirac identify the spikes in the EEG. Root mean instantaneous frequency square (RMIFS) and its parameter ratio from Dirac component are computed and and given to SVM as input for classification. RMIFS is defined as square root over sum of time average squared bandwidth σ_T^2 and center frequency square $< \omega >^2$. These two parameters σ_T^2 extracted from Dirac component and $< \omega >^2$ from LDCT component are also used for classification. These two sets of features shows considerable high accuracy and sensitivity comparable with other existing works

This work is organized as; MCA followed by dataset uses and SVM in Section 2, Materials and Methods. Section 3, Results and Discussion. We end with Section 4, Conclusion.

5.2 Method and Material

In the subsequent subsections, we elaborate MCA first, followed by features computed from its output decomposition. Briefly explained the SVM, the material and data used in this work.

5.2.1 Morphological Component Analysis

Morphological component analysis uses the concept of sparsity and independent redundant transforms to decompose an EEG signal by adapting to the prevailing types of morphologies simultaneously. Representing EEG as sparse linear contribution of coefficients, morphological component analysis (MCA) uses overcomplete dictionary $\Phi \in R^{n \times k}$, where k is the morphological component of an EEG signal $S \in R^n$ decomposed by constructing source component $\{\phi_k\}_{k \in \Gamma}$, where Γ representing the type of explicit dictionaries. An EEG signal can be represented as a sparse linear combination of coefficient. Over-complete dictionary Φ is a set of explicit dictionary, which are defined by a set of mathematical functions to represent the specific morphologies of EEG, Chen et al. [85]. Signal can be represented as

$$\begin{aligned} S &= \textstyle\sum_{i=0}^{k} \beta_i \phi_i + \zeta \\ &= \beta_1\phi_1 + \beta_2\phi_2 \cdots + \beta_k\phi_k + \zeta \\ &\cong s_1 + s_2 \cdots + s_k \ (\zeta \ll 1) \\ &= S' \end{aligned} \tag{5.1}$$

where ϕ_k represents set of basis elements and β is target coefficients to reconstruct the original EEG signal. ζ is assumed to be negligible noise tend to zero. By using three dictionaries undecimated wavelet transform (UDWT), local discrete cosine transform (LDCT) and Dirac (Kronecker basis) [45] [86] [87] in this work, coefficients are optimised as

$$\begin{aligned} &\{\beta_0^{opt}, \beta_1^{opt}, \beta_2^{opt}\} = \underset{\beta_0,\cdots,\beta_k}{\arg\min} \textstyle\sum_{i=0}^{k} \|\beta_i\|_0 \\ &\text{subject to: } S = \textstyle\sum_{i=0}^{k} \beta_i \phi_i \end{aligned} \tag{5.2}$$

The basis pursuit solution [88] was used to represent the sparse component which describe the equation (5.1) as

$$\{\beta_0^{opt}, \beta_1^{opt}, \beta_2^{opt}\} = \underset{\beta_0,\beta_1,\beta_2}{\arg\min} \sum_{i=0}^{2} \|\beta_i\|_1 + \lambda \left\| S - \sum_{i=0}^{2} \phi_i \beta_i \right\|_2^2 \tag{5.3}$$

Equation (5.3) is optimised by block coordinate relaxation (BCR) method [89] in finite time. The algorithm given in [45] as follows:

Data: EEG signal
initiate, number of iteration I_{max} and
threshold: $\delta = \lambda * I_{max}.$;
while $\delta > \lambda$ **do**
  for $p = 0; p \leq 2; p++$ **do**
    Update s_p, considering s_q and s_r fixed.
    1. q=(p+1) mod 3;
       r=(p+2) mod 3;
    2. $R = S - s_q - s_r$;
    3. $\beta_p = \phi_p^T R$;
    4. Threshold the coefficient of β_p
       and obtain $\widehat{\beta}_p$;
    5. Reconstruct s_p by $s_p = \phi_p \widehat{\beta}_p$;
  end
  Update the threshold by $\delta = \delta - \lambda$;
end

Algorithm 2: Block-Coordinate-Relaxation algorithm.

The number of iteration $I_{max} = 100$ is used. Balbir et al. [45] had varied the value of λ from $3 - 5$ depending on type of hard and soft threshold. In this work $\lambda = 3$ is used. Figure 5-1 depicts the working of MCA as described Algorithm 2.

5.2.2 Features Computation

Hilbert transform was applied on the components produced by MCA. Signal is normalised, instantaneous frequency $\varphi'(t)$ and amplitude $a(t)$ were calculated. Following

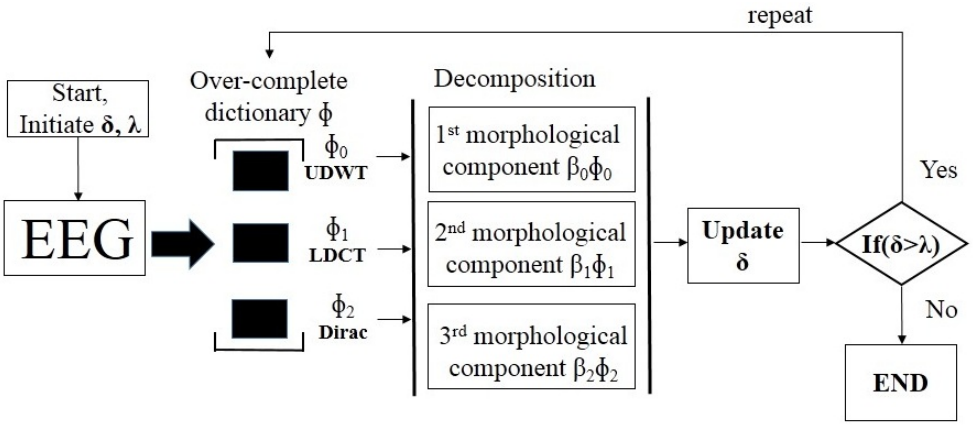


Figure 5-1: Block diagram of MCA decomposition of EEG signal.

the same methodology, by applying hermitian time frequency operator $\left(\frac{1}{j}\frac{d}{dt}\right)$ to equation (3.4). Center frequency is calculated [62].

$$< \omega >= \int \varphi'(t)a^2(t)dt. \tag{5.4}$$

Imaginary part ignored as zero. $a^2(t)$ is density in time [61]. All integrals computed are between time interval $[0, 23.6]$.

The feature combinations been used with MCA are contribution ratio of bandwidth square and RMS frequency:

$$\tau_B = \frac{B_{AM}^2}{B_{FM}^2}, \tag{5.5}$$

$$\Gamma_{RMS} = \frac{B^2}{< \omega >^2}. \tag{5.6}$$

Second combination of features are fractional contribution to dominant frequency

$$\delta_{DF_{<\omega^2>}} = \frac{< \omega^2 >}{f_d}, \tag{5.7}$$

$$\delta_{DF_{AM}} = \frac{B_{AM}^2}{f_d}. \tag{5.8}$$

Third pair of features used are RMIFS frequency and its parameters ratio :

$$f_R = \sqrt{\sigma_T^2 + < \omega >^2}, \tag{5.9}$$

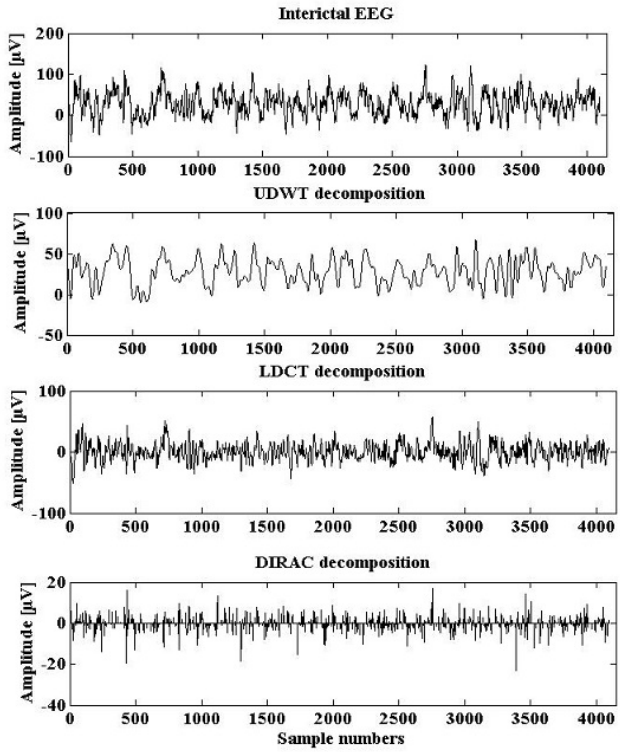


Figure 5-2: MCA decomposition of non seizure interictal EEG signal.

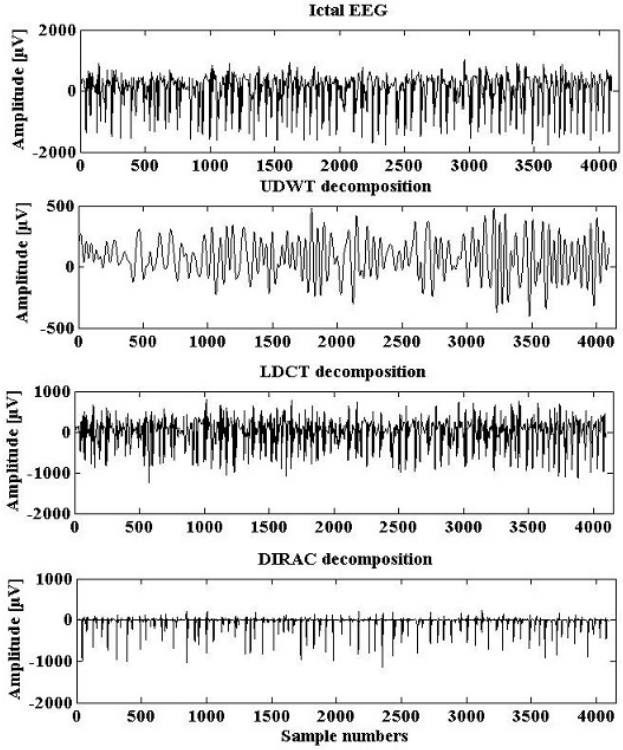


Figure 5-3: MCA decomposition of ictal EEG signal.

$$E_{RMIFS} = \frac{< \omega >^2}{\sigma_T^2}. \tag{5.10}$$

And fourth set of features are set of parameters of RMIFS frequency i.e. $< \omega >^2$ and σ_T^2. Features from RMIFS frequency are depicted in Figure 5-4.

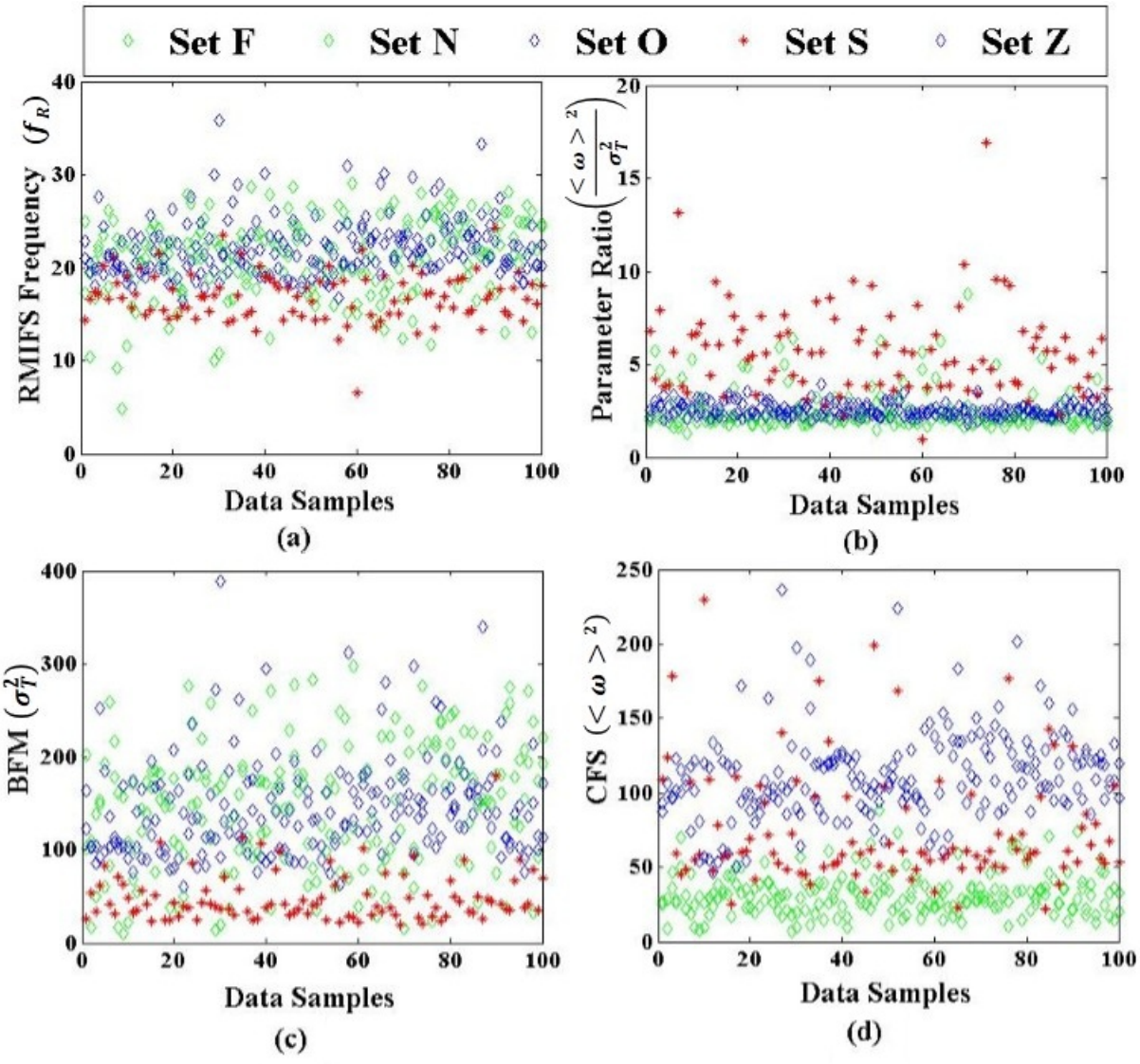


Figure 5-4: Features (a) RMIFS frequency from Dirac component, (b) Parameter ratio of RMIFS frequency from Dirac component (c) Bandwidth amplitude modulation (BFM) from Dirac component (d) Center frequency square (CFS) from LDCT component.

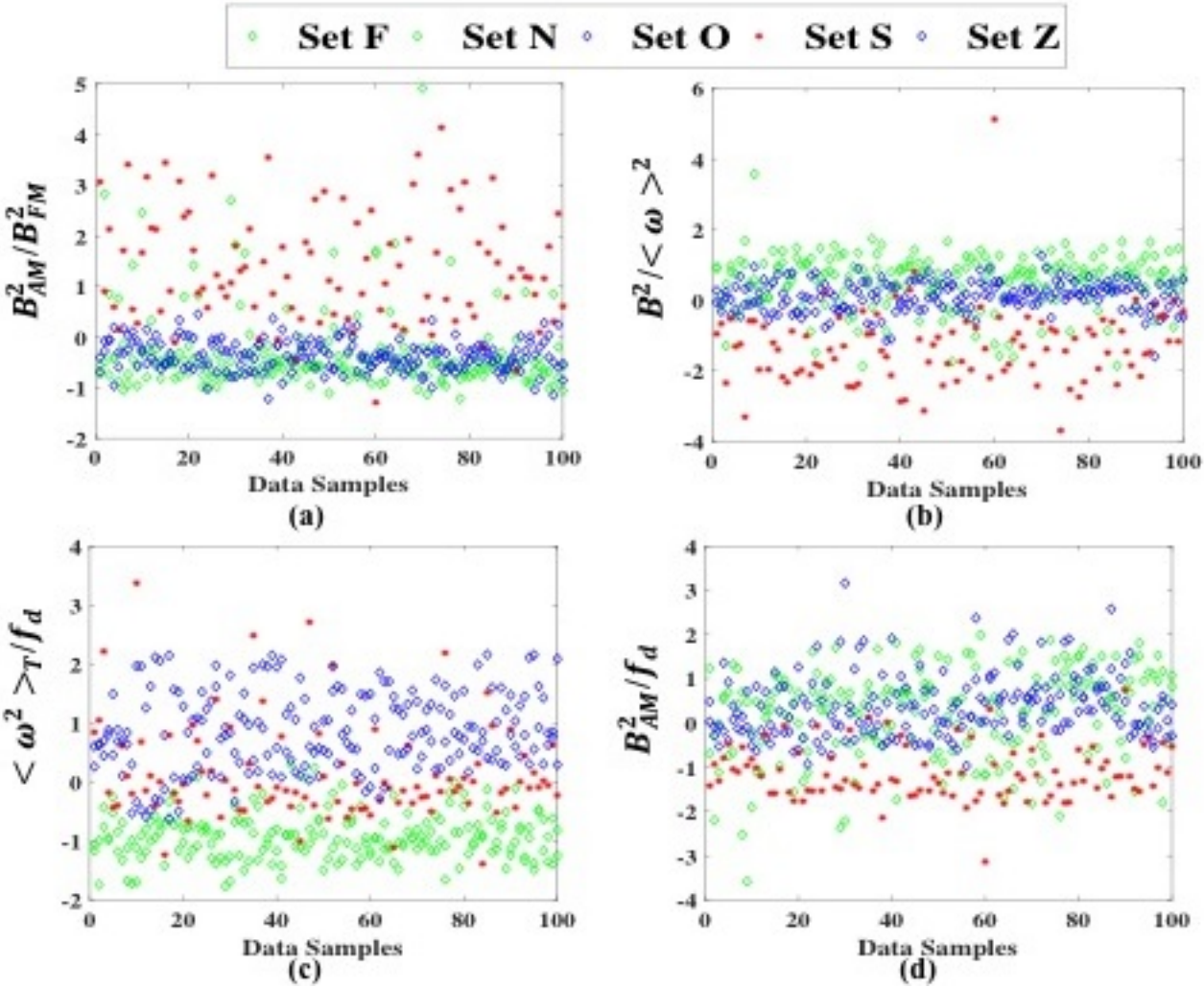


Figure 5-5: Normalized features (a) Parameters ratio of bandwidth square $\frac{B^2_{AM}}{B^2_{FM}}$ (b) parameters ratio of root mean square frequency $\frac{B^2}{<\omega>^2}$ (c) Spectral component of fractional contribution of dominant frequency $\frac{<\omega^2>_T}{f_d}$ (d) Bandwidth component of fractional contribution to dominant frequency.

5.2.3 SVM

Support vector machine (SVM), introduced by Vapnik [64] is used as classifier. SVM discriminate two different classes by creating a hyperplane the maximises distance between among them. Radial Basis function (RBF) kernel is used in this work represented :

$$G(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(\frac{||\boldsymbol{x_i} - \boldsymbol{x_j}||^2}{2\sigma^2}\right), \tag{5.11}$$

where σ is a positive number.

5.2.4 Dataset

EEG dataset [65], Bonn dataset was used to apply the method. This time, six combination of subsets are created for classification. First one with sets F, S, second with N, S and third with O against S, fourth Z against S, fifth F, N versus S and sixth O, Z against S for classification.

Every subset contains 100 data from each feature calculated upon 100 signal segments. These data are normalized using standard deviation and mean. Training and test set are prepared randomly in 70-30 ratio for SVM as previously. Using grid search, kernel parameters are optimised. But in most of the case, default kernel parameters has shown good results as presented in Table 5.2.

5.3 Results and Discussion

In MCA methodology, sparsity play vital role in separating component having different time-frequency properties or morphology of constructing of individual source components. The combination of explicit dictionaries form an over-complete dictionary makes the MCA more powerful methods for de noising and source component separation. Mostly decomposition-based methods required the required the prior knowledge about the decompose components. MCA based decomposition has advan-

tage in accurate reconstruction of the original component because the source component has low probability of occurrence simultaneously. This method relies on the sparsity and over-completeness of the dictionary ($\Phi \in R^{n\times k}$), a set of k redundant transforms, which represent the specific morphologies of different components of signal. Due to the concept of sparsity and over-completeness dictionary extended the traditional signal decomposition to feature extractions of multiple types of morphology simultaneously. As we know that the each construction source signal have specific morphology therefore the EEG time series data can be decomposed by one explicit dictionary and cannot be decomposed by other explicit dictionaries. It is estimated the accurate components as the decomposed components are sparse and independent. The S is the linear combination of different brain activity, where β is the brain activity and Φ is mixing matrix.

After using MCA for decomposition, Hilbert transform was applied over the components which take the real value signal to complex time frequency domain. Feature contribution ratio of bandwidth square and RMS frequency, second fractional contribution of dominant frequency, third RMIFS (f_R) and its parameter ratio ($\frac{<\omega>^2}{\sigma_T^2}$) are computed from LDCT and Dirac component. Contribution ratio of parameters of square bandwidth and RMS frequency are computed from Dirac component where as fractional contribution of dominant frequency are computed from both LDCT and Dirac component. Feature $\frac{<\omega^2>}{f_d}$ was calculated from LDCT component as it shows better significant difference in the set. And $\frac{B_{AM}^2}{f_d}$ was computed from Dirac component for same reason.These feature can be called relative feature of dominant frequency. They can not be seen as normalised feature as dominant frequency f_d remain different for different signal segment. Similarly, feature $\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$ can be relative features.These features $\frac{B_{AM}^2}{f_d}$, $\frac{B_{AM}^2}{B_{FM}^2}$ *and* $\frac{B^2}{<\omega>^2}$, we believe can overcome the drawback of the feature B_{AM}^2 and B^2 without distorting its physical relevance. RMIFS is defined as root over sum of time averaged bandwidth square σ_T^2 and center frequency square $<\omega>^2$. f_R^2 is always greater than $<\omega>^2$ by σ_T^2. The parameter ratio of RMIFS frequency shows how dominant center frequency square is over time averaged bandwidth square. As bandwidth can been taken as standard deviation and center frequency as

mean, f_R satisfy the definition of root mean square. Value of f_R will be close to center frequency when instantaneous frequencies are close to center frequency. Dirac component was chosen for computation of RMIFS frequency because, first it represents the spike morphology of the EEG and secondly, it shows more significant difference than when f_R is computed from LDCT component. For classification using $< \omega >^2$ and σ_T^2, $< \omega >^2$ is computed from from LDCT component and σ_T^2 from Dirac component simultaneously. Because LDCT represents the frequency component better than Dirac which shows modulation better.

Center frequency $< \omega >$ calculated from LDCT component are in range from higher delta wave to lower alpha wave in interictal sets F, N. Whereas in healthy non seizure sets O, Z, it is between higher theta wave to alpha wave. Center frequency in ictal set S are disperse between lower theta wave to lower beta wave. RMIFS f_R on an average is in beta range for all the subset of the Bonn dataset as presented in Figure 5-4.

These feature are normalized using mean and standard deviation then fed to SVM in set of two pairs separately to elaborate its significance in classification of seizures. These pair of features are selected as they are showing opposite behaviour which helps SVM creating hyperplane discriminating the classes. Performance of the SVM classifier is evaluated by using the statistical parameters as previously i.e. sensitivity (SEN), specificity (SPE) and Accuracy (Acc) [20, 36]. where true positive and true negative events are denoted by TP, TN i.e. detecting ictal and interictal correctly. FN and FP stands for false negative, false positive.

Classification of seizure against interictal and healthy non seizure data using bandwidth feature $\frac{B_{AM}^2}{f_d}$, $\frac{B_{AM}^2}{B_{FM}^2}$ shows overall better result than feature fractional contribution of dominant frequency $\frac{<\omega^2>}{f_d}$, $\frac{B_{AM}^2}{f_d}$. In classification of set F against S, 97.65%, 96.75 was observed using bandwidth feature and fractional contribution feature. Set N versus S is the only case where $\frac{<\omega^2>}{f_d}$, $\frac{B_{AM}^2}{f_d}$ performed 99.56% better than $\frac{B_{AM}^2}{f_d}$, $\frac{B_{AM}^2}{B_{FM}^2}$ at 97.65. For set O against S, bandwidth features has classification accuracy of 99.46% where as fractional contribution of dominant frequency has 87.03%. For set Z versus S, bandwidth feature has average accuracy of 100% and fractional contribution feature has 88.15% which was observed with set F, N against S also. Here,

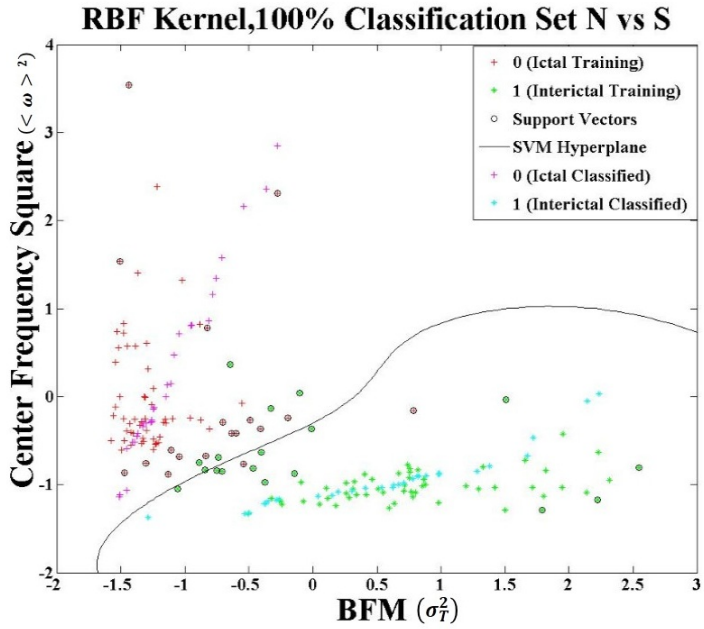


Figure 5-6: Set N vs S classification using RBF kernel $\sigma = 1.0, c = 1.0$.

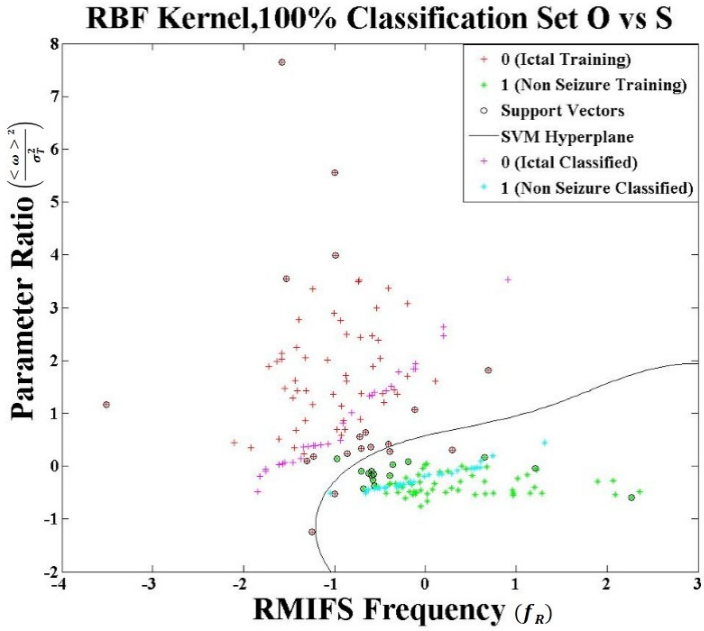


Figure 5-7: Set O vs S classification using RBF kernel $\sigma = 1.0, c = 1.0$.

bandwidth feature has shown slightly low result of 94.83% of accuracy. For set O, Z versus S, bandwidth feature has obtain average classification accuracy of 98.96% whereas fractional contribution has 87.85%. Classification of result of set F versus set S using both pairs of feature i.e. $f_R,\ \frac{<\omega>^2}{\sigma_T^2}$ and $\sigma_T^2,\ <\omega>^2$ shows similar result of average accuracy of 96.48%, 97.13%, average sensitivity of 93.53%, 94.26%. Average specificity using both the features are very high at 99.43%, 100.0%. Results are shown in Table 5.2.

Classification accuracy of both the pairs of feature for set N against S are good at 99.41%, 99.48%. Average sensitivity and specificity are 99.46%, 99.90%, 99.36% and 99.66%. For set O versus Z, features $f_R,\ \frac{<\omega>^2}{\sigma_T^2}$ shows average accuracy of 99.91% but $\sigma_T^2,\ <\omega>^2$ achieved lowly at 87.98%. Similar accuracy result are observed for set Z versus S with 99.63% using $f_R,\ \frac{<\omega>^2}{\sigma_T^2}$ whereas 90.30% using $\sigma_T^2,\ <\omega>^2$. SVM plot for set N versus S using $\sigma_T^2,\ <\omega>^2$ and set O against S using $f_R,\ \frac{<\omega>^2}{\sigma_T^2}$ are shown in Figure 5-6-5-7. Average accuracy for set F, N together versus set S are at 93.05% and 93.61% for both set of features whereas average classification accuracy for set O, Z versus S are at 99.11% and 90.60%. We have compared this proposed work with previous work at Table 5.3-5.4.

For most of time default kernel parameters proved to be better. Even with optimised parameter that are found with grid search are close to default setting and shows little improvement of at most 1%. Therefore, the cases where we found improvement less than 1%, default setting or default kernel parameters were used which helped in avoiding computing overload of parameters search and makes the application more practical. Both the pairs of features has shown similar classification result for interictal set versus ictal or seizure set where as feature $f_R, \frac{<\omega>^2}{\sigma_T^2}$ has shown better results for healthy non seizure classification against seizure set. Therefore, $f_R, \frac{<\omega>^2}{\sigma_T^2}$ features combination for classification are found to better than $\sigma_T^2,\ <\omega>^2$. Figure 5-4 clearly shows it is hard to have 2D map helping SVM to create hyperplane to separate non seizure and seizure sets using feature combination of $\sigma_T^2,\ <\omega>^2$ as they are quite intermingled. Although classification of seizure set which is result of intracranial experiment against non invasive extracranial set is inappropriate yet clas-

sification is been done for comparison purpose of the proposed method with previous works. Detail comparison are presented in Table 5.3-5.4.

5.4 Summary

MCA gives definite number of decomposition depending on the number of set of basis used in over-complete dictionary. This dictionary can be formed based on problem requirements. Selection of basis functions in the dictionary plays important role in creating problem specific application. We found LDCT component is best suited for spectral feature extraction whereas Dirac bases are good in showing spike morphology of the EEG. Default setting of SVM kernel are suitable for proposed features combinations which makes it suitable for practical application. In future, we will try to form a dictionary to remove high oscillation artefacts using MCA with suitable dictionary.

Table 5.1: Classification results from 100 trials on all six combination of sets with contribution ratio of parameters of bandwidth square and RMS frequency, another with fractional contribution of dominant frequency.

	RBF Kernel	SPE [%]	SEN [%]	Acc [%]
Feature	Parameters	min-max avg	min-max avg	min-max avg
Set F vs S				
$\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$	$\sigma = 1.0,\ c = 1.0$	86.66-100.0 98.40	83.33-100.0 96.90	91.66-100.0 97.65
$\frac{<\omega^2>}{f_d}$, $\frac{B_{AM}^2}{f_d}$	$\sigma = 1.0,\ c = 1.0$	96.66-100.0 99.99	83.33-100.0 93.66	91.66-100.0 96.75
Set N vs S				
$\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$	$\sigma = 1.0,\ c = 1.0$	90.00-100.0 98.13	100.0-100.0 100.0	95.00-100.0 97.65
$\frac{<\omega^2>}{f_d}$, $\frac{B_{AM}^2}{f_d}$	$\sigma = 1.0,\ c = 1.0$	93.33-100.0 99.80	90.00-100.0 99.33	95.00-100.0 99.56
Set O vs S				
$\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$	$\sigma = 1.0,\ c = 1.0$	93.33-100.0 98.93	100.0-100.0 100.0	96.66-100.0 99.46
$\frac{<\omega^2>}{f_d}$, $\frac{B_{AM}^2}{f_d}$	$\sigma = 1.0,\ c = 1.0$	80.00-100.0 93.46	66.66-93.33 80.60	80.00-91.66 87.03
Set Z vs S				
$\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$	$\sigma = 1.0,\ c = 1.0$	100.0-100.0 100.0	100.0-100.0 100.0	100.0-100.0 100.0
$\frac{<\omega^2>}{f_d}$, $\frac{B_{AM}^2}{f_d}$	$\sigma = 1.0,\ c = 1.0$	70.00-100.0 89.99	70.00-100.0 86.36	80.00-98.33 88.15
Set F, N vs S				
$\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$	$\sigma = 1.0,\ c = 1.0$	76.66-100.0 90.56	90.00-100.0 99.10	85.00-100.0 94.83
$\frac{<\omega^2>}{f_d}$, $\frac{B_{AM}^2}{f_d}$	$\sigma = 1.0,\ c = 1.0$	73.33-100.0 89.03	90.00-100.0 86.36	80.00-98.33 88.15
Set O, Z vs S				
$\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$	$\sigma = 1.0,\ c = 1.0$	86.66-100.0 97.93	100.0-100.0 100.0	93.33-100.0 98.96
$\frac{<\omega^2>}{f_d}$, $\frac{B_{AM}^2}{f_d}$	$\sigma = 1.0,\ c = 1.0$	83.33-100.0 95.86	60.00-100.0 79.83	78.33-98.33 87.85

Table 5.2: Classification results from 100 trials on all six combination of sets with RMIFS frequency, its parameter's ratio and its parameters.

Feature	RBF Kernel Parameters	SPE [%] min-max avg	SEN [%] min-max avg	Acc [%] min-max avg
Set F vs S				
$f_R,\ \frac{<\omega>^2}{\sigma_T^2}$	$\sigma = 0.99,\ c = 1.0$	86.66-100.0 99.43	80.00-100.0 93.53	90.00-100.0 96.48
$\sigma_T^2,\ <\omega>^2$	$\sigma = 1.0,\ c = 1.0$	100.0-100.0 100.0	83.33-100.0 94.26	91.66-100.0 97.13
Set N vs S				
$f_R,\ \frac{<\omega>^2}{\sigma_T^2}$	$\sigma = 1.0,\ c = 1.0$	93.33-100.0 99.36	93.33-100.0 99.46	96.66-100.0 99.41
$\sigma_T^2,\ <\omega>^2$	$\sigma = 1.0,\ c = 1.0$	93.33-100.0 99.66	96.66-100.0 99.90	96.66-100.0 99.78
Set O vs S				
$f_R,\ \frac{<\omega>^2}{\sigma_T^2}$	$\sigma = 1.0,\ c = 1.0$	93.33-100.0 99.83	100.0-100.0 100.0	96.66-100.0 99.91
$\sigma_T^2,\ <\omega>^2$	$\sigma = 1.0,\ c = 1.0$	86.66-100.0 94.40	66.66-100.0 81.56	81.66-95.00 87.98
Set Z vs S				
$f_R,\ \frac{<\omega>^2}{\sigma_T^2}$	$\sigma = 1.0,\ c = 1.0$	93.33-100.0 99.26	100.0-100.0 100.0	96.66-100.0 99.63
$\sigma_T^2,\ <\omega>^2$	$\sigma = 1.0,\ c = 1.0$	80.00-100.0 91.20	73.33-100.0 89.40	78.33-98.33 90.30
Set F, N vs S				
$f_R,\ \frac{<\omega>^2}{\sigma_T^2}$	$\sigma = 0.96,\ c = 1.0$	73.33-100.0 90.86	86.66-100.0 95.23	85.00-100.0 93.05
$\sigma_T^2,\ <\omega>^2$	$\sigma = 0.92,\ c = 1.0$	80.00-100.0 90.30	86.66-100.0 96.93	86.66-98.33 93.61
Set O, Z vs S				
$f_R,\ \frac{<\omega>^2}{\sigma_T^2}$	$\sigma = 0.85,\ c = 3.0$	90.00-100.0 98.23	100.0-100.0 100.0	95.00-100.0 99.11
$\sigma_T^2,\ <\omega>^2$	$\sigma = 0.77,\ c = 8.0$	86.66-100.0 96.23	70.00-100.0 84.96	81.66-100.0 90.60

Table 5.3: Comparison of classification result of interictal set F, N against ictal set S with other existing works on Bonn dataset.

Author	Decomposition or Preprocessing Method	Features Used	Classifier	Set	Accuracy
Liang et al. [20]	FFT	16 spectral features	SVM	F vs S	98.74%
Nicolau et al. [25]	NA	Permutation entropy	SVM	F vs S	83.13%
				N vs S	88.83%
Siuly et al. [74]	Clustering	9 temporal features	LS-SVM	F vs S	93.91%
				N vs S	97.69%
Zhu et al. [26]	NA	strength and degree of HVG features	K-NN	F vs S	93.00%
				N vs S	98.00%
Riaz et al. [34]	EMD	6 temporal and spectral features	Decision trees	F vs S	96.00%
			SVM	F vs S	93.00%
Samiee et al. [75]	Rational DSTFT	5 time frequency features	MLP	F vs S	94.90%
				N vs S	98.50%
Hassan et al. [30]	CEEMDAN	6 spectral features	Boosting	F vs S	97.00%
				N vs S	100.0%
			SVM	F vs S	93.00%
				N vs S	99.00%
Proposed work	**MCA**	$\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$ σ_T^2, $<\omega>^2$	**SVM**	**F vs S**	**97.65**%
			SVM	**N vs S**	**99.78**%
Sharma et al. [37]	EMD	2D, 3D PSR	LS-SVM	F, N vs S	98.67%
Altunay et al. [76]	L. P Filter	Energy based feature	Threshold	F, N vs S	94.00%
Joshi et al. [77]	FLP	FLP Energy signal energy	SVM	F, N vs S	95.33%
Pachori et al. [78]	EMD	SODP of IMF	ANN	F, N vs S	97.75%
Proposed work	**MCA**	$\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$	**SVM**	**F, N vs S**	**94.83**%

Note: Not applicable (NA), rational discrete STFT (RDSTFT), complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), phase space representation (PSR), linear prediction filter (L. P Filter), fractional linear predict -ion (FLP).

Table 5.4: Comparison with other works on Bonn dataset for classification between healthy non seizure set O, Z and seizure or ictal set S.

Author	**Decomposition or Preprocessing Method**	**Features Used**	**Classifier**	**Set**	**Accuracy**
Guo et al. [15]	Genetic Algorithm	Curve length, standard deviation	KNN	Z vs S	99.20%
Siuly et al. [74]	Clustering	9 temporal features	LS-SVM	Z vs S	99.90%
				O vs S	96.30%
Zhu et al. [26]	NA	strength and degree of HVG features	KNN	Z vs S	100.0%
				O, S	97.00%
Samiee et al. [75]	Rational DSTFT	5 time frequency	MLP	Z vs S	99.80%
Hassan et al. [30]	CEEMDAN	6 spectral features	Boosting	Z vs S	100.0%
Rincon et al. [90]	Wavelet transform	Bag of words (BoW)	SVM	Z vs S	99.85%
		Wavelet coeffiecient	SVM	Z vs S	100.0%
Proposed work	**MCA**	$\frac{B_{AM}^2}{B_{FM}^2}$, $\frac{B^2}{<\omega>^2}$	**SVM**	**Z vs S**	**100.0%**
		f_R, $\frac{<\omega>^2}{\sigma_T^2}$	**SVM**	**O vs S**	**99.91%**
Chen et al. [11]	DTCWT	Logarithm of FFT spectra	NN	Z, O vs S	100%
Proposed work	**MCA**	**RMIFS frequency f_R and its parameters ratio $\frac{<\omega>^2}{\sigma_T^2}$**	**SVM**	**Z, O vs S**	**99.11%**

Chapter 6

Conclusion

The RMS frequency and dominant frequency, we proposed and used in this research in time frequency domain overcomes the drawback of the bandwidth square and instantaneous bandwidth square features which may be identical for two signals. To address these drawback is very important as these features are commonly used and can be interpreted in various ways giving rise to different quantities. Another contribution in this work is utilising the relationship of these two quantities to create a novel feature RMIFS frequency. RMIFS frequency is also average measure and in line with RMS and dominant frequency. The advantage of RMIFS frequency is that it is totally depending phase function and altogether avoiding amplitude based component which can be easily induced with noise and aforementioned drawback.All the proposed and used features are represented in Table 6.1.

As the focus of this work was on feature extraction, we have used basic EMD which is good in handling and decomposing non stationary EEGs and SVM for classification. We have found root mean square frequency f_r or dominant frequency f_d is good average measure in classifying ictal from interictal and non seizure healthy EEG with highest average accuracy of 99.91%. These results are obtained from features of IMF2 having range in alpha, beta brain wave. Average sensitivity of 100% was observed using RMS frequency. Mean instantaneous frequency square in terms of time averaged instantaneous frequency spread and center frequency is also good in discriminating ictal from interictal with 98.71% of average accuracy. We found selection and combi-

Table 6.1: Proposed time frequency features.

Feature	Expression
RMS frequency	$f_r = \sqrt{< \omega >_T^2 + B_{AM}^2 + B_{FM}^2}.$
Fractional contribution to RMS frequency	$\gamma_{RMS_{AM}} = \frac{B_{AM}^2}{f_r},$ $\gamma_{RMS_{FM}} = \frac{B_{FM}^2}{f_r},$ $\gamma_{RMS_{<\omega>_T^2}} = \frac{< \omega >_T^2}{f_r}.$
Parameters ratio of RMS frequency	$\Gamma_{RMS} = \frac{< \omega >_T^2}{B^2}.$
Dominant frequency	$f_d = \sqrt{B_{AM}^2 + < \omega^2 >_T}.$
Fractional contribution to dominant frequency	$\delta_{DF_{AM}} = \frac{B_{AM}^2}{f_d},$ $\delta_{DF_{<\omega^2>_T}} = \frac{< \omega^2 >_T}{f_d},$
Parameters ratio of dominant frequency	$\Delta_{DF} = \frac{< \omega^2 >_T}{B_{AM}^2}.$
RMIFS frequency	$f_R = \sqrt{B_{FM}^2 + < \omega >_T^2}.$
Fractional contribution to RMIFS frequency	$\varepsilon_{RMIFS_{FM}} = \frac{B_{FM}^2}{f_R},$ $\varepsilon_{RMIFS_{<\omega>_T^2}} = \frac{< \omega >_T^2}{f_R},$
Parameters ratio of RMIFS frequency	$E_{RMIFS} = \frac{< \omega >_T^2}{B_{FM}^2}.$

nation of features displaying opposite behaviours gives good results. All the features used in this work exhibit best and consistent results with default setting of SVM. This makes them suitable for practical clinical trial as it reduce the computational overload of optimization of SVM kernel parameters. In future we will try to find better feature to address the problem of false positive using our proposed adaptive thresholding method.Adaptive thresholding method has scope of improvisation. Deciding on to which IMF will best classification remains a problem which increases the computational overhead. We had addressed this issue with MCA, a new decomposition method. We had utilized rest of the features like fractional contribution of dominant frequency and contribution ratio of RMS frequency and square bandwidth along with RMIFS frequency for classification. MCA has an advantage of reconstructing the signal with minimum coefficients and gives freedom to choose bases functions for dictionary depending on the application. MCA was used to address the indefinite number of IMFs problem which increases the computational overhead in extracting the features. Every methods has some advantage and disadvantages and yet to be complete in itself. EMD decomposes the EEGs adaptively giving indefinite number of non orthogonal IMFs where as MCA gives definite number decomposition depending on the dictionary and reconstruct the signal with minimum number of coefficients but finally is represented as linear combination of transforms which gives independent orthogonal decomposition by considering the signal weakly stationary. To show the combination of features plays good role in classification result, we kept the feature set with EMD constant and varying the feature set with MCA to take the best result as presented in the Figure 6-1. All these feature are time frequency based features and can be used with any time series data that shows amplitude modulation and frequency modulation like geo-acoustic, seismic, ECG, EMG and radar signal classification for example.

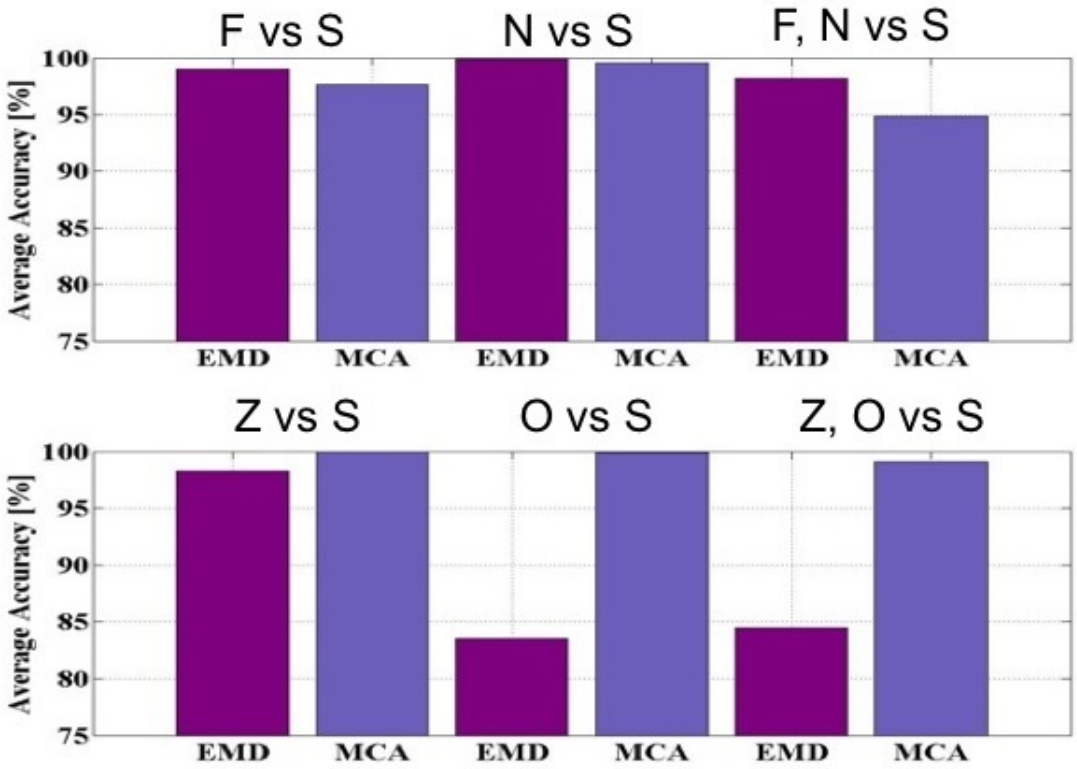


Figure 6-1: Comparison of EMD and MCA based on classification results on Bonn dataset.

Bibliography

[1] Anne T Berg, Samuel F Berkovic, Martin J Brodie, Jeffrey Buchhalter, J Helen Cross, Walter van Emde Boas, Jerome Engel, Jacqueline French, Tracy A Glauser, Gary W Mathern, et al. Revised terminology and concepts for organization of seizures and epilepsies: report of the ilae commission on classification and terminology, 2005–2009. *Epilepsia*, 51(4):676–685, 2010.

[2] Robert S Fisher, Walter van Emde Boas, Warren Blume, Christian Elger, Pierre Genton, Phillip Lee, and Jerome Engel. Epileptic seizures and epilepsy: definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe). *Epilepsia*, 46(4):470–472, 2005.

[3] V Srinivasan, C Eswaran, Sriraam, and N. Artificial neural network based epileptic detection using time-domain and frequency-domain features. *Journal of Medical Systems*, 29(6):647–660, 2005.

[4] Kemal Polat and Salih Güneş. Classification of epileptiform eeg using a hybrid system based on decision tree classifier and fast fourier transform. *Applied Mathematics and Computation*, 187(2):1017–1026, 2007.

[5] Ivan Osorio, Mark G Frei, Jon Giftakis, Tom Peters, Jeff Ingram, Mary Turnbull, Michele Herzog, Mark T Rise, Scott Schaffner, Richard A Wennberg, et al. Performance reassessment of a real-time seizure-detection algorithm on long ecog series. *Epilepsia*, 43(12):1522–1535, 2002.

[6] Alexandros T Tzallas, Markos G Tsipouras, and Dimitrios I Fotiadis. Epileptic seizure detection in eegs using time–frequency analysis. *IEEE transactions on information technology in biomedicine*, 13(5):703–710, 2009.

[7] Hojjat Adeli, Ziqin Zhou, and Nahid Dadmehr. Analysis of eeg records in an epileptic patient using wavelet transform. *Journal of neuroscience methods*, 123(1):69–87, 2003.

[8] James S Walker. *A primer on wavelets and their scientific applications*. CRC press, 2008.

[9] Anne Simon Moffat, Elizabeth J Maggio, David Holzman, Michelle Hoffman, TA Heppenheimer, Addison Greenwood, Andrew Chaikin, Barbara Burke, Marcia F Bartusiak, et al. *A positron named Priscilla: scientific discovery at the frontier*. National Academies Press, 1994.

[10] Hasan Ocak. Optimal classification of epileptic seizures in eeg using wavelet analysis and genetic algorithm. *Signal processing*, 88(7):1858–1867, 2008.

[11] Guangyi Chen. Automatic eeg seizure detection using dual-tree complex wavelet-fourier features. *Expert Systems with Applications*, 41(5):2391–2394, 2014.

[12] YU Khan and J Gotman. Wavelet based automatic seizure detection in intracerebral electroencephalogram. *Clinical Neurophysiology*, 114(5):898–908, 2003.

[13] Scott B Wilson, Mark L Scheuer, Ronald G Emerson, and Andrew J Gabor. Seizure detection: evaluation of the reveal algorithm. *Clinical neurophysiology*, 115(10):2280–2291, 2004.

[14] ME Saab and Jean Gotman. A system to detect the onset of epileptic seizures in scalp eeg. *Clinical Neurophysiology*, 116(2):427–442, 2005.

[15] Ling Guo, Daniel Rivero, Julián Dorado, Cristian R Munteanu, and Alejandro Pazos. Automatic feature extraction using genetic programming: an application

to epileptic eeg classification. *Expert Systems with Applications*, 38(8):10425–10436, 2011.

[16] Abdulhamit Subasi and M Ismail Gursoy. Eeg signal classification using pca, ica, lda and support vector machines. *Expert Systems with Applications*, 37(12):8659–8666, 2010.

[17] Katrien Jansen and Lieven Lagae. Cardiac changes in epilepsy. *Seizure*, 19(8):455–460, 2010.

[18] K Lehnertz and CE Elger. Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neuronal complexity loss. *Electroencephalography and clinical Neurophysiology*, 95(2):108–117, 1995.

[19] Agostino Accardo, M Affinito, M Carrozzi, and F Bouquet. Use of the fractal dimension for the analysis of electroencephalographic time series. *Biological cybernetics*, 77(5):339–350, 1997.

[20] Sheng-Fu Liang, Hsu-Chuan Wang, and Wan-Lin Chang. Combination of eeg complexity and spectral analysis for epilepsy diagnosis and seizure detection. *EURASIP Journal on Advances in Signal Processing*, 2010(1):853434, 2010.

[21] Shengkun Xie, Feng Jin, and Sridhar Krishnan. Sparse approximation of long-term biomedical signals for classification via dynamic pca. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 7167–7170. IEEE, 2011.

[22] Zafer Iscan, Zümray Dokur, and Tamer Demiralp. Classification of electroencephalogram signals with combined time and frequency features. *Expert Systems with Applications*, 38(8):10499–10505, 2011.

[23] Y Tang and DM Durand. A tunable support vector machine assembly classifier for epileptic seizure detection. *Expert systems with applications*, 39(4):3925–3938, 2012.

[24] A Temko, E Thomas, W Marnane, G Lightbody, and G Boylan. Eeg-based neonatal seizure detection with support vector machines. *Clinical Neurophysiology*, 122(3):464–473, 2011.

[25] Nicoletta Nicolaou and Julius Georgiou. Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. *Expert Systems with Applications*, 39(1):202–209, 2012.

[26] Guohun Zhu, Yan Li, and Peng Paul Wen. Epileptic seizure detection in eegs signals using a fast weighted horizontal visibility algorithm. *Computer methods and programs in biomedicine*, 115(2):64–75, 2014.

[27] U Rajendra Acharya, S Vinitha Sree, and Jasjit S Suri. Automatic detection of epileptic eeg signals using higher order cumulant features. *International journal of neural systems*, 21(05):403–414, 2011.

[28] WRS Webber, Ronald P Lesser, Russell T Richardson, and Kerry Wilson. An approach to seizure detection using an artificial neural network (ann). *Electroencephalography and clinical Neurophysiology*, 98(4):250–272, 1996.

[29] Bruno Direito, César Teixeira, Bernardete Ribeiro, Miguel Castelo-Branco, Francisco Sales, and António Dourado. Modeling epileptic brain states using eeg spectral analysis and topographic mapping. *Journal of neuroscience methods*, 210(2):220–229, 2012.

[30] Ahnaf Rashik Hassan and Abdulhamit Subasi. Automatic identification of epileptic seizures from eeg signals using linear programming boosting. *computer methods and programs in biomedicine*, 136:65–77, 2016.

[31] Norden E Huang, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society of London A: mathemat-*

ical, physical and engineering sciences, volume 454, pages 903–995. The Royal Society, 1998.

[32] Rami J Oweis and Enas W Abdulhay. Seizure classification in eeg signals utilizing hilbert-huang transform. *Biomedical engineering online*, 10(1):38, 2011.

[33] Shufang Li, Weidong Zhou, Qi Yuan, Shujuan Geng, and Dongmei Cai. Feature extraction and recognition of ictal eeg using emd and svm. *Computers in biology and medicine*, 43(7):807–816, 2013.

[34] Farhan Riaz, Ali Hassan, Saad Rehman, Imran Khan Niazi, and Kim Dremstrup. Emd-based temporal and spectral features for the classification of eeg signals using supervised learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(1):28–35, 2016.

[35] Kai Fu, Jianfeng Qu, Yi Chai, and Yong Dong. Classification of seizure based on the time-frequency image of eeg signals using hht and svm. *Biomedical Signal Processing and Control*, 13:15–22, 2014.

[36] Varun Bajaj and Ram Bilas Pachori. Classification of seizure and nonseizure eeg signals using empirical mode decomposition. *IEEE Transactions on Information Technology in Biomedicine*, 16(6):1135–1142, 2012.

[37] Rajeev Sharma and Ram Bilas Pachori. Classification of epileptic seizures in eeg signals based on phase space representation of intrinsic mode functions. *Expert Systems with Applications*, 42(3):1106–1117, 2015.

[38] Arindam G Mahapatra and Keiichi Horio. Overcoming drawback of feature instantaneous bandwidth using emd for epileptic seizure classification by rms frequency. In *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on*, pages 001322–001327. IEEE, 2016.

[39] Katarzyna J Blinowska and Piotr J Durka. Unbiased high resolution method of eeg analysis in time-frequency space. *Acta Neurobiologiae Experimentalis*, 61(3):157–174, 2001.

[40] Elaheh Imani, Hamid-Reza Pourreza, and Touka Banaee. Fully automated diabetic retinopathy screening using morphological component analysis. *Computerized Medical Imaging and Graphics*, 43:78–88, 2015.

[41] Elaheh Imani, Malihe Javidi, and Hamid-Reza Pourreza. Improvement of retinal blood vessel detection using morphological component analysis. *Computer methods and programs in biomedicine*, 118(3):263–279, 2015.

[42] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5):411–430, 2000.

[43] Patrick Berg and Michael Scherg. A multiple source approach to the correction of eye artifacts. *Electroencephalography and clinical neurophysiology*, 90(3):229–241, 1994.

[44] Garrick L Wallstrom, Robert E Kass, Anita Miller, Jeffrey F Cohn, and Nathan A Fox. Automatic correction of ocular artifacts in the eeg: a comparison of regression-based and component-based methods. *International journal of psychophysiology*, 53(2):105–119, 2004.

[45] Balbir Singh and Hiroaki Wagatsuma. A removal of eye movement and blink artifacts from eeg data using morphological component analysis. *Computational and mathematical methods in medicine*, 2017, 2017.

[46] Kaveh Samiee, Péter Kovács, and Moncef Gabbouj. Epileptic seizure detection in long-term eeg records using sparse rational decomposition and local gabor binary patterns feature extraction. *Knowledge-Based Systems*, 118:228–240, 2017.

[47] Jiří Spilka, Jordan Frecon, Roberto Leonarduzzi, Nelly Pustelnik, Patrice Abry, and Muriel Doret. Sparse support vector machine for intrapartum fetal heart rate classification. *IEEE journal of biomedical and health informatics*, 21(3):664–671, 2017.

[48] Fengyu Cong, Qiu-Hua Lin, Li-Dan Kuang, Xiao-Feng Gong, Piia Astikainen, and Tapani Ristaniemi. Tensor decomposition of eeg signals: a brief review. *Journal of neuroscience methods*, 248:59–69, 2015.

[49] Mohammad Zavid Parvez and Manoranjan Paul. Epileptic seizure detection by analyzing eeg signals using different transformation techniques. *Neurocomputing*, 145:190–200, 2014.

[50] Ali A Abdul-latif, Irena Cosic, Dinesh K Kumar, Barbara Polus, and Cliff Da Costa. Power changes of eeg signals associated with muscle fatigue: the root mean square analysis of eeg bands. In *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004. Proceedings of the 2004*, pages 531–534. IEEE, 2004.

[51] Piyush Swami, Tapan K Gandhi, Bijaya K Panigrahi, Manjari Tripathi, and Sneh Anand. A novel robust diagnostic model to detect seizures in electroencephalography. *Expert Systems with Applications*, 56:116–130, 2016.

[52] William C Stacey and Brian Litt. Technology insight: neuroengineering and epilepsy?designing devices for seizure control. *Nature clinical practice Neurology*, 4(4):190–201, 2008.

[53] Sriram Ramgopal, Sigride Thome-Souza, Michele Jackson, Navah Ester Kadish, Iván Sánchez Fernández, Jacquelyn Klehm, William Bosl, Claus Reinsberger, Steven Schachter, and Tobias Loddenkemper. Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy. *Epilepsy & behavior*, 37:291–307, 2014.

[54] Patrick J Loughlin and Berkant Tacer. Comments on the interpretation of instantaneous frequency. *IEEE Signal Processing Letters*, 4(5):123–125, 1997.

[55] B Boashash, M Mesbah, and P Coldtiz. Time-frequency detection of eeg abnormalities, chapter 15, article 15.5 in [8], 2003.

[56] Gabriel Rilling and Patrick Flandrin. On the influence of sampling on the empirical mode decomposition. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 3, pages III–III. IEEE, 2006.

[57] Gabriel Rilling and Patrick Flandrin. One or two frequencies? the empirical mode decomposition answers. *IEEE transactions on signal processing*, 56(1):85–95, 2008.

[58] L Mandel. Interpretation of instantaneous frequencies. *American Journal of Physics*, 42(10):840–846, 1974.

[59] Athanasios Papoulis and S Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

[60] Arthur E Barnes. Instantaneous spectral bandwidth and dominant frequency with applications to seismic reflection data. *Geophysics*, 58(3):419–428, 1993.

[61] Leon Cohen and Chongmoon Lee. Instantaneous bandwidth for signals and spectrogram. In *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, pages 2451–2454. IEEE, 1990.

[62] Leon Cohen. *Time-frequency analysis*, volume 778. Prentice Hall PTR Englewood Cliffs, NJ:, 1995.

[63] Leon Cohen and Chongmoon Lee. Instantaneous frequency, its standard deviation and multicomponent signals. In *Proc. SPIE*, volume 975, pages l86–208, 1988.

[64] Vladimir Vapnik. *The nature of statistical learning theory*. Springer science & business media, 2013.

[65] Ralph G Andrzejak, Klaus Lehnertz, Florian Mormann, Christoph Rieke, Peter David, and Christian E Elger. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, 64(6):061907, 2001.

[66] Shufang Li, Weidong Zhou, Qi Yuan, Shujuan Geng, and Dongmei Cai. Feature extraction and recognition of ictal eeg using emd and svm. *Computers in biology and medicine*, 43(7):807–816, 2013.

[67] Kai Fu, Jianfeng Qu, Yi Chai, and Yong Dong. Classification of seizure based on the time-frequency image of eeg signals using hht and svm. *Biomedical Signal Processing and Control*, 13:15–22, 2014.

[68] Naveed Rehman and Danilo P Mandic. Multivariate empirical mode decomposition. In *Proceedings of The Royal Society of London A: Mathematical, Physical and Engineering Sciences*, page rspa20090502. The Royal Society, 2009.

[69] Patrick J Loughlin and Keith L Davidson. Instantaneous kurtosis. *IEEE Signal Processing Letters*, 7(6):156–159, 2000.

[70] Susan Tolwinski. The hilbert transform and empirical mode decomposition as tools for data analysis. *Tucson: University of Arizona*, 2007.

[71] Junsheng Cheng, Dejie Yu, Jiashi Tang, and Yu Yang. Application of frequency family separation method based upon emd and local hilbert energy spectrum method to gear fault diagnosis. *Mechanism and Machine Theory*, 43(6):712–723, 2008.

[72] Patrick J Loughlin and Keith L Davidson. Instantaneous spectral skew and kurtosis. In *Statistical Signal and Array Processing, 2000. Proceedings of the Tenth IEEE Workshop on*, pages 574–578. IEEE, 2000.

[73] Keith L Davidson and Patrick J Loughlin. Instantaneous spectral moments. *Journal of the Franklin Institute*, 337(4):421–436, 2000.

[74] Yan Li, Peng Paul Wen, et al. Clustering technique-based least square support vector machine for eeg signal classification. *Computer methods and programs in biomedicine*, 104(3):358–372, 2011.

[75] Kaveh Samiee, Peter Kovacs, and Moncef Gabbouj. Epileptic seizure classification of eeg time-series using rational discrete short-time fourier transform. *IEEE transactions on Biomedical Engineering*, 62(2):541–552, 2015.

[76] Semih Altunay, Ziya Telatar, and Osman Erogul. Epileptic eeg detection using the linear prediction error energy. *Expert Systems with Applications*, 37(8):5661–5665, 2010.

[77] Varun Joshi, Ram Bilas Pachori, and Antony Vijesh. Classification of ictal and seizure-free eeg signals using fractional linear prediction. *Biomedical Signal Processing and Control*, 9:1–5, 2014.

[78] Varun Joshi, Ram Bilas Pachori, and Antony Vijesh. Classification of ictal and seizure-free eeg signals using fractional linear prediction. *Biomedical Signal Processing and Control*, 9:1–5, 2014.

[79] Ira J Rampil. A primer for eeg signal processing in anesthesia. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 89(4):980–1002, 1998.

[80] Benoît Crépon, Vincent Navarro, Dominique Hasboun, Stéphane Clemenceau, Jacques Martinerie, Michel Baulac, Claude Adam, and Michel Le Van Quyen. Mapping interictal oscillations greater than 200 hz recorded with intracranial macroelectrodes in human epilepsy. *Brain*, 133(1):33–45, 2009.

[81] He Sheng Liu, Tong Zhang, and Fu Sheng Yang. A multistage, multimethod approach for automatic detection and classification of epileptiform eeg. *IEEE Transactions on Biomedical Engineering*, 49(12):1557–1566, 2002.

[82] Katarzyna J Blinowska and Piotr J Durka. Unbiased high resolution method of eeg analysis in time-frequency space. *Acta Neurobiologiae Experimentalis*, 61(3):157–174, 2001.

[83] Elaheh Imani, Hamid-Reza Pourreza, and Touka Banaee. Fully automated diabetic retinopathy screening using morphological component analysis. *Computerized Medical Imaging and Graphics*, 43:78–88, 2015.

[84] Elaheh Imani, Malihe Javidi, and Hamid-Reza Pourreza. Improvement of retinal blood vessel detection using morphological component analysis. *Computer methods and programs in biomedicine*, 118(3):263–279, 2015.

[85] Scott Shaobing Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM review*, 43(1):129–159, 2001.

[86] Markus Püschel and José MF Moura. The algebraic approach to the discrete cosine and sine transforms and their fast algorithms. *SIAM Journal on Computing*, 32(5):1280–1316, 2003.

[87] Xuancheng Shao and Steven G Johnson. Type-iv dct, dst, and mdct algorithms with reduced numbers of arithmetic operations. *Signal Processing*, 88(6):1313–1326, 2008.

[88] Jean-Luc Starck, Y Moudden, J Bobin, M Elad, and DL Donoho. Morphological component analysis. In *Proceedings of SPIE*, volume 5914, pages 1–15, 2005.

[89] S Sardy, AG Bruce, and P Tseng. Block coordinate relaxation methods for nonparametric signal denoising with wavelet dictionaries. 1998.

[90] Jesus Martinez-del Rincon, Maria J Santofimia, Xavier del Toro, Jesus Barba, Francisca Romero, Patricia Navas, and Juan C Lopez. Non-linear classifiers applied to eeg analysis for epilepsy seizure detection. *Expert Systems with Applications*, 2017.