Emotion Recognition Based on Statistical Modeling of EMD-DWT Transformed EEG Signals Responsive to Music Videos

by

S. M. Shafiu1 Hasan

MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING
Department of Electrical and Electronic Engineering

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY

02 January, 2019
The thesis entitled “Emotion Recognition Based on Statistical Modeling of EMD-DWT Transformed EEG Signals Responsive to Music Videos” submitted by Student Name: S. M. Shafiqui Hasan, Student No.: 0416062222P, Session: April, 2016 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING on 02 January, 2018.

BOARD OF EXAMINERS

1. (Dr. Celia Shahnaz)  
   Professor  
   Department of Electrical and Electronic Engineering  
   Bangladesh University of Engineering and Technology  
   Dhaka - 1000, Bangladesh.  
   Chairman  
   (Supervisor)

2. (Dr. Md. Shafiqui Islam)  
   Professor and Head  
   Department of Electrical and Electronic Engineering  
   Bangladesh University of Engineering and Technology  
   Dhaka - 1000, Bangladesh.  
   Member  
   (Ex-officio)

3. (Dr. Mohammed Imamul Hassan Bhuiyan)  
   Professor  
   Department of Electrical and Electronic Engineering  
   Bangladesh University of Engineering and Technology  
   Dhaka - 1000, Bangladesh.  
   Member

4. (Dr. Mohammed Moshiul Hoque)  
   Professor  
   Department of Electrical and Electronic Engineering  
   Chittagong University of Engineering and Technology  
   Chittagong, Bangladesh.  
   Member  
   (External)
CANDIDATE'S DECLARATION

I, do, hereby declare that neither this thesis nor any part of it has been submitted elsewhere for the award of any degree or diploma.

Signature of the Candidate

[Signature]

S. M. Shahidul Hasan
Dedication

To my loving family.
Acknowledgment

I would like to express my sincere gratitude to my supervisor Dr. Celia Shahnaz for her guidance, encouragement, words of advise and continuous moral support during the span of the research. I also want to thank her for her valuable time in counseling me about various aspects of not only research but also life itself. I am greatly indebted to her for her sincere efforts to build up the ability of constructive thinking in me which I believe has helped me grow as a researcher.

I would also like to thank the rest of the members of my thesis committee: Prof. Dr. Md. Shafiqul Islam, Dr. Mohammed Imamul Hassan Bhuiyan, and Dr. Mohammed Moshiul Hoque, for their encouragement and insightful comments. I would like to thank the head of the department of Electrical and Electronic Engineering for allowing me to use the lab facilities, which contributed greatly in completing the work in time. I would also like to thank Professor Dr. Shaikh Anowarul Fattah for his informative suggestions and kind words of encouragement.

Special note of thanks to my wife Samira Naushin Alam for her continuous moral support, and heartily cooperation throughout this journey. Thanks to my younger brother S. M. Toufiquil Hasan for helping me to start fresh everyday. Last but not the least, and most importantly, I wish to thank my parents, who raised me, supported me and taught me with utmost patience and love.
Abstract

In the recent years, implementation of brain-computer interface (BCI) for implicit emotion tagging of multimedia contents has become a very popular area of research. In this thesis, a statistical modeling based emotion recognition method is proposed by utilizing EMD-DWT transformed Electroencephalogram (EEG) signals responsive to music videos. By applying Empirical Mode Decomposition (EMD), a set of dominant Intrinsic Mode Functions (IMFs) are selected on which a Discrete Wavelet Transform (DWT) is performed. Then various statistical models are fitted to the transformed DWT coefficients to find out the appropriate models which describe the coefficients well based on Bayesian Information Criterion. The parameters of the selected models are used to form the feature vector. Furthermore, an efficient feature selection method is formulated using ReliefF algorithm to avoid redundant features obtained from the feature extraction scheme. The reduced feature set thus obtained is then fed to a Support Vector Machine (SVM) to perform two class classification of different dimensions describing emotions. Extensive simulations are carried out to test the efficacy of the proposed method using DEAP, an affective computing database. It is found that the proposed method outperforms some state-of-the-art methods in terms of accuracy and F1-score.
Contents

Dedication iii

Acknowledgements iv

Abstract v

1 INTRODUCTION 1
1.1 Introduction ......................................................... 1
1.2 Definition of Emotion ............................................... 1
1.3 Characteristics of EEG Signals ................................. 2
  1.3.1 Electroencephalography and brain waves ................. 2
  1.3.2 Location of EEG electrodes ................................ 4
  1.3.3 Paradigms of EEG ............................................ 5
1.4 Reasons for Using EEG for Emotion Recognition .......... 6
1.5 Problem Statement .................................................. 7
1.6 Motivation of the Proposed Approach ......................... 7
1.7 Objectives and Outcome of the Thesis ....................... 8
1.8 Organization of the Thesis ....................................... 8

2 LITERATURE REVIEW 9
2.1 Introduction ......................................................... 9
2.2 Emotion Description Model .................................. 9
2.3 EEG Signal Preprocessing .................................. 10
2.4 Existing EEG-based Emotion Recognition Methods .......... 12
2.5 Existing Feature Extraction Methods ....................... 14
  2.5.1 Types of EEG features .................................. 14
  2.5.2 Feature extraction process ............................... 15
2.6 Available Feature and Channel Selection Methods in EEG-based Emotion Recognition ........................................... 15
2.7 Existing Classification Methods .............................................. 16
   2.7.1 Classifiers ................................................................. 16
   2.7.2 Offline vs online ....................................................... 16
   2.7.3 User dependence ...................................................... 17
2.8 Conclusion ...................................................................... 17

3 A Method of Emotion Recognition Based on Statistical Modeling of EMD- Wavelet Transformed EEG Signals Responsive to Music Videos .................................................. 18
3.1 Introduction ..................................................................... 18
3.2 Outline of the Proposed Method ......................................... 18
3.3 Selection of EEG Channels from Neurological Viewpoint ............. 20
3.4 EMD-based EEG Signal Decomposition .................................. 21
3.5 Proposed Selection Method of Dominant IMFs ....................... 23
3.6 Discrete Wavelet Transform of the Dominant IMFs .................. 24
3.7 Proposed Statistical Modeling of the DWT coefficients ............... 25
   3.7.1 Definition of statistical modeling ................................... 25
   3.7.2 Statistical modeling according to BIC criterion ................ 26
3.8 Proposed Feature Extraction Process .................................... 28
3.9 Effectiveness of the Proposed Features .................................. 28
   3.9.1 All sub-bands’ feature set ........................................... 28
   3.9.2 $\alpha, \beta$ and $\gamma$ sub bands’ feature set ....................... 39
   3.9.3 Euclidean distance between median values of features of different classes ...................................................... 52
   3.9.4 Measure of intra class compactness ................................ 53
3.10 Proposed Feature Selection Method .................................... 54
3.11 Change in Inter class Separability and Intra Class Compactness .... 55
3.12 Discussion on Selected Features ......................................... 56
3.13 Classification Methods ..................................................... 58
   3.13.1 Support Vector Machine (SVM) ................................... 58
   3.13.2 K-nearest Neighbors (KNN) ....................................... 59
   3.13.3 Naive Bayes (NB) ..................................................... 59
3.13.4 Decision Tree ............................................. 60
3.13.5 Ensemble Classifier ........................................ 60
3.14 Conclusion ...................................................... 61

4 SIMULATION RESULTS AND PERFORMANCE EVALUATION 63
  4.1 Introduction .................................................. 63
  4.2 Dataset Description .......................................... 63
  4.3 Data Processing .............................................. 64
  4.4 Simulation Results ........................................... 64
  4.5 Comparison with Existing Works .......................... 69
  4.6 Discussion ................................................... 70
  4.7 Conclusion ................................................... 71

5 CONCLUSION ..................................................... 73
  5.1 Concluding Remarks ......................................... 73
  5.2 Contributions of the thesis ................................. 73
  5.3 Scopes for Future Work ..................................... 74

Bibliography ..................................................... 76
List of Tables

3.1 Different level DWT coefficients and their corresponding sub bands of EEG signal .................................................. 25
3.2 Statistical models and model parameters for feature vector formation .................................................. 28
3.3 Comparison of median of feature values between high and low classes for all subbands’ feature set .......................................... 40
3.4 Comparison of median of feature values between high and low classes for $\alpha, \beta$ and $\gamma$ sub bands’ feature set .................................................. 51
3.5 Comparison of Euclidean distance between median of feature values of high and low classes for all sub bands’ and $\alpha, \beta$ and $\gamma$ sub bands’ feature sets expressed as percentage of range of feature values) .................................................. 52
3.6 Comparison of Euclidean distance between all feature values and median of feature values of high and low classes for all sub bands’ feature set expressed as percentage of range of feature values .................................................. 53
3.7 Comparison of Euclidean distance between all feature values and median of feature values of high and low classes for $\alpha, \beta$ and $\gamma$ sub bands’ feature set expressed as percentage of range of feature values .................................................. 54
3.8 Change in inter class separability and intra class compactness for all sub bands’ feature set after feature selection .................................................. 56
3.9 Change in inter class separability and intra class compactness for $\alpha, \beta$ and $\gamma$ sub bands’ feature set after feature selection .................................................. 56
3.10 Number of features selected using ReliefF in different emotion dimension .................................................. 57
4.1 Classification accuracy and its standard deviation without feature selection .................................................. 65
4.2 Classification accuracy and its standard deviation after feature selection With ReliefF .................................................. 66
4.3 Accuracy comparison With existing works .................................................. 69
4.4 F1-score comparison With existing works . . . . . . . . . . . . 70
List of Figures

1.1 The cortex subdivided into the frontal, temporal, parietal, and occipital lobes. ......................................................... 3
1.2 The five brain waves: $\delta, \theta, \alpha, \beta$, and $\gamma$. ................................................................. 3
1.3 The International 10/20 system. ....................................................... 5

2.1 Valence-Arousal plane and 2-D SAM (Self-Assessment Manikin) questionnaire (Different emotions can be located in the V-A plane by calculating the arousal and valence values based on 2-D SAM questionnaire.) . . . 10

3.1 Block diagram for the proposed method ................................................. 19
3.2 Selected channels for the proposed method; Red denotes the channels selected on the left hemisphere, blue denotes the channels selected in the right hemisphere and green denotes the channels selected in the central area of the brain. ....................................................... 21
3.3 An original EEG signal from a subject and the resulting IMFS. ............. 23
3.4 Summary of the statistical models which best fitted the data according to BIC criterion ....................................................... 26
3.5 Comparison between median of $k$ (shape) parameter of GEV distribution for all sub bands’ coefficient set in high and low valence classes . . . 29
3.6 Comparison between median of $\sigma$ (scale) parameter of GEV distribution for all sub bands’ coefficient set in high and low valence classes . . . 30
3.7 Comparison between median of $\mu$ (location) parameter of GEV distribution for all sub bands’ coefficient set in high and low valence classes . . . 30
3.8 Comparison between median of $\mu$ (mean) parameter of Normal distribution for all sub bands’ coefficient set in high and low valence classes . . . 31
3.9 Comparison between median of $\sigma$ (standard deviation) parameter of Normal distribution for all sub bands' coefficient set in high and low valence classes .................................................. 31
3.10 Comparison between median of $k$ (shape) parameter of GEV distribution for all sub bands' coefficient set in high and low arousal classes .......... 32
3.11 Comparison between median of $\sigma$ (scale) parameter of GEV distribution for all sub bands' coefficient set in high and low arousal classes .......... 32
3.12 Comparison between median of $\mu$ (location) parameter of GEV distribution for all sub bands' coefficient set in high and low arousal classes ........ 33
3.13 Comparison between median of $\mu$ (mean) parameter of Normal distribution for all sub bands' coefficient set in high and low arousal classes ........ 33
3.14 Comparison between median of $\sigma$ (standard deviation) parameter of Normal distribution for all sub bands' coefficient set in high and low arousal classes .................................................. 34
3.15 Comparison between median of $k$ (shape) parameter of GEV distribution for all sub bands' coefficient set in high and low dominance classes ...... 34
3.16 Comparison between median of $\sigma$ (scale) parameter of GEV distribution for all sub bands' coefficient set in high and low dominance classes ...... 35
3.17 Comparison between median of $\mu$ (location) parameter of GEV distribution for all sub bands' coefficient set in high and low dominance classes .... 35
3.18 Comparison between median of $\mu$ (mean) parameter of Normal distribution for all sub bands' coefficient set in high and low dominance classes .... 36
3.19 Comparison between median of $\sigma$ (standard deviation) parameter of Normal distribution for all sub bands' coefficient set in high and low dominance classes .................................................. 36
3.20 Comparison between median of $k$ (shape) parameter of GEV distribution for all sub bands' coefficient set in high and low liking classes ............ 37
3.21 Comparison between median of $\sigma$ (scale) parameter of GEV distribution for all sub bands' coefficient set in high and low liking classes ............ 37
3.22 Comparison between median of $\mu$ (location) parameter of GEV distribution for all sub bands' coefficient set in high and low liking classes ............ 38
3.23 Comparison between median of μ (mean) parameter of Normal distribution for all sub bands’ coefficient set in high and low liking classes . 38
3.24 Comparison between median of σ (standard deviation) parameter of Normal distribution for all sub bands’ coefficient set in high and low liking classes . 39
3.25 Comparison between median of μ (location) parameter of EV distribution for α, β and γ sub bands’ coefficient set in high and low valence classes . 41
3.26 Comparison between median of σ (scale) parameter of EV distribution for α, β and γ sub bands’ coefficient set in high and low valence classes . 41
3.27 Comparison between median of k (shape) parameter of GEV distribution for α, β and γ sub bands’ coefficient set in high and low valence classes . 42
3.28 Comparison between median of σ (scale) parameter of GEV distribution for α, β and γ sub bands’ coefficient set in high and low valence classes . 42
3.29 Comparison between median of μ (location) parameter of GEV distribution for α, β and γ sub bands’ coefficient set in high and low valence classes . 43
3.30 Comparison between median of μ (location) parameter of EV distribution for α, β and γ sub bands’ coefficient set in high and low arousal classes . 43
3.31 Comparison between median of σ (scale) parameter of EV distribution for α, β and γ sub bands’ coefficient set in high and low arousal classes . 44
3.32 Comparison between median of k (shape) parameter of GEV distribution for α, β and γ sub bands’ coefficient set in high and low arousal classes . 44
3.33 Comparison between median of σ (scale) parameter of GEV distribution for α, β and γ sub bands’ coefficient set in high and low arousal classes . 45
3.34 Comparison between median of μ (location) parameter of GEV distribution for α, β and γ sub bands’ coefficient set in high and low arousal classes . 45
3.35 Comparison between median of μ (location) parameter of EV distribution for α, β and γ sub bands’ coefficient set in high and low dominance classes 46
3.36 Comparison between median of σ (scale) parameter of EV distribution for α, β and γ sub bands’ coefficient set in high and low dominance classes 46
3.37 Comparison between median of k (shape) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands' coefficient set in high and low dominance classes

3.38 Comparison between median of $\sigma$ (scale) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands' coefficient set in high and low dominance classes

3.39 Comparison between median of $\mu$ (location) parameter of GEV distribution in high and low dominance classes

3.40 Comparison between median of $\mu$ (location) parameter of EV distribution for $\alpha, \beta$ and $\gamma$ sub bands' coefficient set in high and low liking classes

3.41 Comparison between median of $\sigma$ (scale) parameter of EV distribution for $\alpha, \beta$ and $\gamma$ sub bands' coefficient set in high and low liking classes

3.42 Comparison between median of k (shape) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands' coefficient set in high and low liking classes

3.43 Comparison between median of $\sigma$ (scale) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands' coefficient set in high and low liking classes

3.44 Comparison between median of $\mu$ (location) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands' coefficient set in high and low liking classes

3.45 Weights of features obtained from DWT coefficients corresponding to all sub bands

3.46 Weights of features obtained from DWT coefficients corresponding to all sub bands

4.1 Classification accuracy and its standard deviation without feature selection

4.2 Classification accuracy and its standard deviation after feature selection With ReliefF

4.3 Participant wise classification accuracy in valence recognition with mean and standard deviation marked in red

4.4 Participant wise classification accuracy in arousal recognition with mean and standard deviation marked in red

4.5 Participant wise classification accuracy in dominance Recognition with mean and standard deviation marked in red

4.6 Participant wise classification accuracy in liking recognition with mean and standard deviation marked in red
List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMD</td>
<td>Empirical Mode Decomposition</td>
</tr>
<tr>
<td>DWT</td>
<td>Discrete Wavelet Transform</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>KNN</td>
<td>K-Nearest Neighbors</td>
</tr>
<tr>
<td>DWT</td>
<td>Discrete Wavelet Transform</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalogram</td>
</tr>
<tr>
<td>HCI</td>
<td>Human Computer Interaction</td>
</tr>
<tr>
<td>DEAP</td>
<td>Database for Emotion Analysis using Physiological and Audiovisual Signal</td>
</tr>
<tr>
<td>ERNN</td>
<td>Emotion recognition Neural Network</td>
</tr>
<tr>
<td>SAM</td>
<td>Self Assessment Manikin</td>
</tr>
<tr>
<td>DT-CWPT</td>
<td>Dual Tree Continuous Wavelet Packet transform</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular Value Decomposition</td>
</tr>
<tr>
<td>STFT</td>
<td>Short Time Fourier Transform</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform</td>
</tr>
<tr>
<td>PSD</td>
<td>Power Spectral Density</td>
</tr>
<tr>
<td>AE</td>
<td>Approximate Entropy</td>
</tr>
<tr>
<td>DE</td>
<td>Differential Entropy</td>
</tr>
<tr>
<td>HOC</td>
<td>Higher Order Crossings</td>
</tr>
<tr>
<td>CSP</td>
<td>Common Spatial Patterns</td>
</tr>
<tr>
<td>HFD</td>
<td>Higuchi Fractal Dimension</td>
</tr>
<tr>
<td>AI</td>
<td>Asymmetry Index</td>
</tr>
<tr>
<td>IMF</td>
<td>Intrinsic Mode Function</td>
</tr>
<tr>
<td>EV</td>
<td>Extreme value</td>
</tr>
<tr>
<td>GEV</td>
<td>Generalized Extreme value</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

1.1 Introduction

Emotion is a part and parcel of human cognition. It is a subjective phenomenon which can become apparent through conscious or unconscious physiological, biological or mental reactions or expressions [1]. In the advent of progressive brain-computer interface and human-machine interaction, mutual communication between humans and machines has become an issue of great importance. For efficient interfacing, the machines are needed to be equipped with the measures to explain and perceive emotions of human being. Besides, affective characteristics are important features for describing various multimedia contents. The study of human emotions should therefore be considered a very important factor in order to design and implement an effective and efficient Human-Computer Interface. However, emotions have been largely ignored in this particular field. Over the last decade, affects computing has emerged to fulfill this gap by converging technology and emotions into HCI. The aim of affective computing is to study the emotional state of a user and use the information to model emotional interactions between a human and a computer [2]. Since a sustainable improvement can be brought about in automatic recommendation system for multimedia contents and also in efficient Human-Computer Interfacing systems, gaining precise knowledge of users emotions has become a significant research area.

1.2 Definition of Emotion

An emotion is a complex psychological state that involves three distinct components: a subjective experience, a physiological response, and a behavioral or expressive response. Emotions have been described as discrete and consistent responses to events (external or
internal) with significance for the organism. They are brief in duration and correspond to a coordinated set of responses, which may include verbal, behavioral, physiological and neural mechanisms.

In affective neuroscience, the emotion concept can be differentiated from similar constructs like feelings, moods and affects. Feelings can be viewed as a subjective representation of emotions. Moods are diffuse affective states that generally last for much longer duration than emotions and are also usually less intense than emotions. Finally, affect is an encompassing term, used to describe the topics of emotions, feelings, and moods all together.

1.3 Characteristics of EEG Signals

In the following Subsections, we shortly introduce the main characteristics of the EEG signals, to give some context to the reader.

1.3.1 Electroencephalography and brain waves

The largest portion of the human brain, the cortex, is divided into the frontal, temporal, parietal, and occipital lobes (See Figure 1.1) [3]. The frontal lobe is responsible for the conscious thought. The temporal lobe is responsible for the senses of smell and sound, and the processing of complex stimuli such as faces and scenes. The parietal lobe is responsible for integrating sensory information from various senses, as well as the manipulation of objects. Finally, the occipital lobe is responsible for the sense of sight.
Fig. 1.1: The cortex subdivided into the frontal, temporal, parietal, and occipital lobes.

Fig. 1.2: The five brain waves: $\delta$, $\theta$, $\alpha$, $\beta$, and $\gamma$. 
EEG measures voltage fluctuations resulting from ionic current flows within the neurons of the brain and generally a typical adult scalp EEG signal is in the range pf about 10-100 μV [4]. These signals observed in the scalp are divided into specific frequency ranges that are more prominent in certain states of mind. Namely the components are: δ (1-4 Hz), θ (4-8 Hz), α (8-16 Hz), β (16-32 Hz), and γ (greater than 32 Hz) bands [5] (see Fig. 1.2). The beginning and the end of the bands varies a few Hertz among different authors.

Delta waves are associated with the unconscious mind, and occur during a deep dreamless sleep. Theta brain waves are associated with the subconscious mind, for instance with activities such as sleeping and dreaming. Alpha waves are typically associated to a relaxed mental state, yet aware, and are more visible over the parietal and occipital lobes. High alpha activity has been correlated to brain inactivation. Beta waves are related to an active state of mind, more prominent in the frontal cortex and over other areas during intense focused mental activity. Finally, gamma waves are associated with an hyper brain activity.

1.3.2 Location of EEG electrodes

In order to produce replicable setups, there are standardized sets of locations for electrodes on the skull, such as the International 10/20 System (IS) (see Fig. 1.3). This system is based on the relationship between the location of an electrode and the underlying area of the cerebral cortex.

The numbers 10 and 20 indicate the distance between adjacent electrodes (10% or 20% of the total front-back or right-left distance of the skull). Extra positions can be added by the utilization of the existing empty spaces. Each site has a letter to identify the lobe and a number to identify the hemisphere location. F stands for Frontal, T for Temporal, C for Central (although there is no central lobe, C letter is used for identification purposes), P for Parietal, and O for Occipital. z (zero) refer to an electrode placed on the mid line. Even numbers refer to electrode positions on the right hemisphere, while odd numbers refers to the left one. Four anatomical landmarks are used for the correct positioning of the electrodes: nasion (the point between the forehead and nose), inion (the lowest point of the skull from the back of the head, indicated by a prominent bump), and the pre auricular points anterior to the ear. Electrodes can be monopolar or bipolar. The first record the
potential difference, compared to a neutral electrode connected to an ear lobe or mastoid. The second shows the potential difference between two paired electrodes. With the use of high-density electrodes, multiple sources of noise that can disrupt EEG recordings arise, such as muscle activity near the active sites, eye movements and blinks. Eye movement artifacts can have profound effects on frontal brain sites, specifically mid-frontal sites (F3 and F4), commonly used in studying emotional reactivity.

1.3.3 Paradigms of EEG

In order to understand how the changes that occur in the electrical brain activity can be evaluated, we present the paradigms most commonly used: Sensory Evoked Potentials (SEP), Event-Related Potentials (ERP), and Event-Related De/Synchronizations (ERD/ERS).

An evoked potential corresponds to an electrical potential signal recorded after the presentation of a stimulus. There are three types: Auditory Evoked Potentials (AEP), Visual Evoked Potentials (VEP), and the Somatosensory Evoked Potentials (SsEP), that differ by the elicitation method used. AEP are elicited by a click or tone stimulus presented through earphones, VEP by a flashing light or changing pattern on a monitor (Steady State
Visually Evoked Potential (SSVEP) if it is elicited by a periodic stimulus, and SsEP by electrical stimulation of the peripheral nerve.

ERP have a very high temporal resolution that allows the measurement of immediate responses to short stimuli. They are usually measured as latencies and amplitudes of positive and negative potentials at specific millisecond intervals following a stimulus. The ERP components can be encapsulated in the following order: P100, N100, N200, P200, P300, and Slow Cortical Potential (SCP). N100 is characterized by a negative deflection in voltage with a delay between stimulus and response (latency) of 100 ms after the stimulus, while P100 is the equivalent but with a positive deflection. N200 and P200 are analogous to N100 and P100, with a latency of about 200 ms instead of 100 ms (varying between 150 and 275 ms). P300 is thought to reflect processes involved in stimulus evaluation or categorization, and it is characterized by a positive deflection in voltage with a latency of roughly 250 to 500 ms. SCP can occur from 300 ms to over several seconds [6].

ERD/ERS analysis allows for the evaluation of power changes within specified frequency bands with a high temporal resolution. They measure rapid changes of power within defined frequency band ranges in order to assess responses that occur within milliseconds of a stimulus presentation. The increased power within a frequency band after the presentation of a stimulus is defined as an ERS, while ERD corresponds to the decrease of the power within a frequency band. It is appropriate for measuring the existing reactions to affective communications as they occur [6].

1.4 Reasons for Using EEG for Emotion Recognition

There are a number of ways to understand a persons inner emotional state. The major ways are: subjective experiences or feelings, internal or physiological signals, and external or audio/visual signals [7]. Subjective self-reports can be informative about a person's feelings. But this should not be considered as a reliable and valid metric to measure a person's emotional state. That is mostly due to a possibility of contradiction and not understanding the parameters to be measured [8]. There is a chance that the response of the participants may portray how they feel others would answer rather than showing exactly how they are feeling. Due to this fact, researchers tend to investigate the physiological responses responsive to emotional triggers. As physiological signals can assist in obtaining a better understanding of the participants underlying responses expressed at the
time of the observations without any chance of self-conscious contamination, they have become a very popular tool in the field of emotion recognition. More specifically, the noninvasive, fast, and inexpensive nature of EEG, has established it as a preferred method in studying the brains responses to emotional stimuli [9]. Nowadays, wearable, low-cost, easy-to-use, wireless EEG devices are emerging in the market, making it possible to use EEG-based emotion recognition in different areas such as entertainment, e-learning, virtual worlds, or e-health care applications [10], [11]. As a result it has opened the door to a vast area of application and a very promising field of research.

1.5 Problem Statement

The existing problems in the area of EEG-based emotion recognition are manifold. Few of the major problems are stated in this Section.

- Firstly, most of the works carried out on emotion recognition using EEG ignores the value of using neuro-physiology in feature extraction and selection and fails to connect the state of interest of the work to neuro-physiology.

- Secondly, as emotions are subjective, and their recognition depend on their intensity evoked by the stimuli, establishment of proper ground truth is a complex task.

- In addition, special care is required to successfully extract discriminating features in time, frequency or time-frequency domain from EEG signals because of their non-stationery and non-linear nature.

- Moreover, statistical modeling of the features and its effectiveness on classification performance has not been yet explored and still remains a daunting challenge.

1.6 Motivation of the Proposed Approach

To overcome the problems stated in the previous Section, this thesis presents a novel emotion recognition scheme which utilizes the selection of proper EEG channels and frequency ranges to recognize emotions in reference with established psychological and neurological ground truths. This ensures the extraction of informative EEG features for robust and accurate classification of the associated emotions. Additionally, in order to
effectively retain both time and frequency domain information of non-stationary and non-linear EEG data, an EMD-DWT based method is utilized. Moreover, the best fit statistical models are derived from the empirical mode decomposed and the wavelet transformed EEG signals responsive to music videos, given the fact that a well fitted statistical model can not only help estimate uncertainties in the observed data but also estimate probabilistic future behavior of a system based on past statistical information, thus forming an effective feature set from the selected model parameters.

1.7 Objectives and Outcome of the Thesis

The objectives of this thesis are:

- To select proper EEG channels and frequency ranges to recognize emotion in reference with established physiological and neurological ground truths.

- To derive the best fit statistical models from the empirical mode decomposed and the wavelet transformed EEG signals responsive to music videos thus forming an effective feature set from the model parameters.

- To formulate an efficient feature selection method to avoid redundant features obtained from the feature extraction scheme.

- To analyze the performance of the proposed method for different classifiers in comparison to the state-of-the art methods using the same publicly available dataset.

The outcome of this thesis is the development of an efficient emotion recognition method based on a statistical modeling approach by using EEG signals in EMD-DWT domain and analysis of its effectiveness compared to the state of the art methods.

1.8 Organization of the Thesis

The remainder of this thesis is organized as follows. Chapter 2 introduces the basic EEG emotion recognition theories and reviews the related work. Chapter 3 contains the detailed description of the proposed statistical modeling based method carried out on EMD-Wavelet transformed EEG signals. Chapter 4 realizes the proposed method based on the widely-used DEAP database. Additionally, the comparisons with the existing methods and related work are also shown in this Chapter. Chapter 5 concludes this thesis.
Chapter 2

LITERATURE REVIEW

2.1 Introduction

In this Chapter of the thesis, the existing emotion description models, available emotion recognition schemes and feature and channel selection methods are summarized. In addition to that the merits and demerits of different approaches are briefly discussed. In conclusion, the potential of the proposed statistical modeling based methodology is discussed in terms of enhanced effectiveness and efficiency in accordance with the works available in the literature.

2.2 Emotion Description Model

Historically, two basic models have been used for describing human emotions: discrete model and dimensional model [12]. The discrete model comprises of several basic emotions and all the other emotions are regarded as combinations of these basic emotions. However, as emotion is a complex psychological parameter, the ambiguity and vagueness involved in determining appropriate basic emotions is too high. That is why the dimensional model has become preferable and popular to the researchers in this field. In the dimensional model, a hyperspace is proposed to describe human emotions. Typically, the hyperspace is simplified to a two-dimensional valence-arousal plane. Sometimes this model also includes other auxiliary dimensions such as dominance and liking [12], [13]. Different emotions are located in this plane by calculating the valence and arousal values. Specifically, valence values indicate the polarity of human emotions (positive/negative) while arousal values indicate the level of excitement under different moods (calmness/excitement). The emotion labels are determined by subjects themselves according to the Self-Assessment Manikin (SAM) questionnaire [13], [14]. To be more detailed, when
Fig. 2.1: Valence-Arousal plane and 2-D SAM (Self-Assessment Manikin) questionnaire (Different emotions can be located in the V-A plane by calculating the arousal and valence values based on 2-D SAM questionnaire.)

Valence changes from 1 to 9, the emotional states change from negative to positive emotions. Correspondingly, the level of excitement increases as arousal changes from 1 to 9. Dominance is the metric of the feeling of empowerment, and when this dimension changes from 1 to 9 the person is considered to be feeling more and more empowered or in control. Liking simply conveys the information about the extent to which the user likes the emotional trigger being presented. Some researchers also describe other forms of dimensional model such as Russell Valence-Arousal plane [15]. However, the model as shown in Fig. 2.1 is utilized in this paper to realize the valence-arousal 2D discrimination. Also, the level of dominance and liking have also been determined and classified accordingly.

2.3 EEG Signal Preprocessing

The pre-processing of raw EEG data for further analysis involves the use of digital signal processing techniques. Noting that useful data for emotion recognition are at frequencies
between 4-45 Hz, for efficient preprocessing the data is downsampled to 128 Hz and bandpass filtered with a common bandwidth [16], [17]. In common average reference (CAR) the value of the entire electrode montage is subtracted from the channel of interest resulting in a spatial voltage distribution of mean zero. As it emphasizes components that are present in most electrodes, it reduces such components and thereby functions as a highpass filter [18].

The most challenging task of signal pre-processing is to remove noise from EEG without distorting the signals related to emotions. Noise sources,(e.g., external and environmental noise from the EEG equipment, and electromagnetic (EM) noise), are easy to address by ensuring the equipment is in good working order and removing EM sources from the recording room. A more challenging task is to remove physiological noise, such as due to cardiac signals, movements caused by muscle contraction (electromyogram (EMG)), and signal caused by eyeball movement (electrooculogram (EOG)). Noise due to EMG can be minimised by asking the subjects to sit in a comfortable position, and noise due to EOG can be minimised by using stimuli that do not require eye movement. However, it is not possible to completely avoid the aforementioned noise sources. In addition, using stimuli that do not invoke eye movement is not always practical, as they trigger a strong emotional response. Hence, a compromise between avoiding noise due to EOG and having good emotional stimuli has to be found.

To make noise removal easier, some datasets include additional physiological signals like EOG, EMG, plethysmograph, body temperature and measurement from respiration belt. There are two main approaches to EOG artefact correction, namely regression based methods, and methods based on spatial decomposition.

The regression based methods smooth EEG by regressing out the reference EOG signals, e.g., [19], [20], [21], [22]. These require the recordings of EOG signals. In addition, several muscle groups require different reference channels, which makes this approach impractical [23].

The spatial decomposition based methods include principal component analysis (PCA), singular value decomposition (SVD), and blind source decomposition (BSS) such as independent component analysis (ICA). Similar to regression methods, EOG artefacts need to be known. PCA has been shown to be better in removing EOG artefacts than regression methods [23]. In [23], ICA has been shown to recover more brain activity signals than
PCA, and therefore ICA based methods are more favourable.

In addition, one can use fully automatic methods, where artefacts are removed without any prior knowledge of the signals, e.g., the second order blind identification which involves the ICA.

2.4 Existing EEG-based Emotion Recognition Methods

Emotion is the representation of an intricate psychological response corresponding to an external stimuli and it originates directly from human brain. So, in order to correctly and reliably describe human emotions, a tool is needed which can directly reflect the relationship between external stimuli from physical world and actual inner responses of the human race. The properties of EEG matches perfectly to this requirement because through EEG the electro-physiological changes in human brain in response to any physical word stimulus can be closely monitored. As a result emotion recognition using EEG is irreplaceable in the research roadmap of emotion recognition [24], [25].

In the literature, multi-channel EEG signals are recorded and decoded to achieve good performance based on the typical 10-20 system. This points to the fact that different functional cortexes are connected with EEG emotion sensing and recognition tasks. Chanel et al. selected different emotional stimulation sources namely pictures, recall and games as in [23], [24], [26], respectively and recorded subjects EEG signals and physiological signals. They extracted power spectrum density (PSD) features and utilized naive Bayes (NB), support vector machine (SVM), and linear discriminant analysis (LDA) classification methods to recognize 3 emotional states in the valence-arousal plane and achieved an accuracy ranging from 56-63%.

Alzoubi et al. executed an experiment on 3 subjects and used PSD features and a k-nearest neighbors (KNN) classifier to distinguish self-elicited emotions [27]. Lin et al. analyzed EEG spectral power asymmetry (ASM) features from signals acquired by precise medical grade equipment to recognize 4 states with the music emotional stimulations. Eventually, they reached a relatively high recognition accuracy 82.29% using SVM classifier [28].

To further improve the emotion recognition performance, some researcher looked to improving or developing new classification algorithms. Inspired by biological structure, Khosrowabadi et al. proposed an emotion recognition neural network (ERNN) based on
EmoCog architecture to discriminate emotion from EEG. They evaluated the performance of ERNN employing their own dataset and reached a higher recognition accuracy (66-68%) than common methods like KNN and SVM [25].

Later on, other research groups also developed their own emotion recognition frameworks with different alternative stimulations (e.g., sound, movie and video clip), extracted features (e.g., higher order crossings, common spatial pattern and self-organizing map) and classification methods (e.g., quadratic discriminant analysis, deep learning network and fuzzy support vector machine) [29], [30], [31]. An important contribution came from Koelstra et al. who established a multimodal emotion recognition database named DEAP [13]. DEAP database is a very well-known multimodal database for emotion analysis using EEG, physiological and music video signals. This database was later utilized by many research groups in their works [12], [32], [33], [34], [35], [36], [37], [38], [39]. Since this paper will also utilize DEAP database, related work based on this database will be reviewed here.

Soleymani et al. utilized only EEG signals to obtain the valence, arousal recognition accuracy as 52.4%, 57.0%. They also showed a significant improvement after the fusion of multimodal signals, as the valence, arousal classification accuracy became 67.7%, 76.1% [32]. Jirayucharoensak et al. applied deep learning theories on 32-channel EEG data and prelearned the feature set with two-layer stacked-autoencoder. They obtained a valence, arousal recognition accuracy as 54%, 53% [30]. Daimi and Saha proposed a DT-CWPT based feature extraction method and utilized SVD, QRep and F-ratio to obtain a valence, arousal recognition accuracy of 64.3% and 66.2% [33]. Zhuang X. et al. proposed a compact and unsupervised EEG response representation method for emotion recognition and reached a recognition accuracy of 70.9% and 67.1% in valence and arousal dimension respectively [34]. Shahnaz et al. used an EMD-DWT based approach with higher order statistical features for emotion recognition and reached an accuracy of 64.71%, 66.51%, 66.88%, 70.52% in valence, arousal, dominance and liking dimensions respectively using an SVM-RBF classifier. [35] Yin et al. proposed a three-layer stacked-autoencoder ensemble classifier and by combining multimodal signals with EEG, they obtained the valence, arousal recognition accuracy as 84%, 83% which was the highest among the related work employing DEAP database [36]. Zhunag N. et al. reached an accuracy of 69.1% and 71.99% in valence and arousal domain by using multimodal infor-
mation in the EMD modain [37]. Mert and Akan presented a multivariate empirical mode decomposition (MEMD) based scheme with independent component analysis (ICA) and obtained a valence, arousal recognition accuracy of 72.87% and 75% using ANN. [38] Dai et al. proposed a sparsity constrained differential evolution enabled feature-channel-sample hybrid selection for EEG emotion recognition and achieved a valence, arousal classification accuracy of 74.25% and 75.37%. [39]

Most of the related work utilized EEG signals from all the channels provided by publicly available databases or established their own multi-channel EEG sensing system to acquire the raw data. However, this leads to the difficulty of handling and processing large scale raw EEG data, large computational complexity and greater latency. And also this large number of raw data may comprise of redundant and uncorrelated features and channels which may further hinder reliable signal acquisition and data processing [40]. Consequently, an efficient selection scheme is a must to help reduce the sensing channel and extract relevant information from EEG raw waves for emotion recognition applications. In the following Section we discuss about the available EEG feature and channel selection schemes available in literature.

2.5 Existing Feature Extraction Methods

In the following Subsections, we present the most common features extracted from the EEG signals, as well as the methods used to perform it.

2.5.1 Types of EEG features

Regarding the types of EEG features that authors used, around 10% of the works do not provide any information, while the remaining used mainly the delta, theta, alpha, beta, and gamma bands (89.4%). Almost 37% of these used all the bands together, while the remaining selected only some of them, such as alpha, beta, theta, and gamma (13.7%), alpha and beta (7.8%), alpha, beta, and gamma (7.8%), delta, theta, alpha, and beta (3.92%), alpha, beta, gamma (3.92%), among other combinations.

The remaining features used were the Event-Related De/Synchronizations (ERD/ERS), Event-Related Potentials (ERP), and fixed frequency bandwidths (e.g., 0.5-30 Hz, 1-10 Hz, 1-46 Hz, and 2-30 Hz).
2.5.2 Feature extraction process

The feature extraction process can be handled using various methods. The most used feature extraction methods are the Fourier Transform such as the Short-time Fourier Transform (STFT) or Discrete Fourier Transform (DFT), statistical, Power Spectral Density (PSD), Wavelet Transform (WT), Entropy such as the Approximate Entropy (AE), Differential Entropy (DE), Sample Entropy (SE), or Wavelet Entropy (WE), Higher Order Crossings (HOC), Common Spatial Patterns (CSP), Fractal Dimensions (mainly the Higuchi Fractal Dimension (HFD)), and Asymmetry Index (AI), Empirical Mode Decomposition (EMD) and its several variations [6].

2.6 Available Feature and Channel Selection Methods in EEG-based Emotion Recognition

EEG signals are high dimensional, hence the computational processing of these signals is often complex and expensive. The purpose of feature extraction is to simplify the subsequent emotion classification task by identifying the important elements of the signal, creating a feature vector based on these elements and using the vector to classify the corresponding emotion. In terms of feature selection, Jenke et al. reviewed all the frequently-used features in emotion recognition and compared the performance of different feature selection algorithms, including ReliefF and mRmR (min-Redundancy-Max-Relevance) [41].

For channel selection there are a number of categorized methods available in the literature: classification accuracy ranking-based, statistical feature-based and common spatial pattern-based methodologies. [39] Thomas et al. first ran an experiment evaluating the classification accuracy ranking of 3 basic classifiers and proposed a channel selection scheme according to the obtained classification accuracies [42]. Later on, some statistical features were brought into action for evaluating the importance of various EEG channels. Claudia et al. selected 17 key channels out of 199 channels in total by a combined evaluation of the P-value and the accuracy ranking results [43]. The common spatial pattern can represent the temporal-spatial connection between different channels. Therefore, Peng et al., Mahnaz et al. and Colwell et al. proposed their optimized channel selection methods with different sparsity parameters or termed jump-wise regression [40], [44], [45].
2.7 Existing Classification Methods

In the field of recognition of emotions we have a large number of classifiers families that are commonly used: bayesian, support vector machines, decision trees, among others. In the following Subsections, we present the most used classifiers, the type of classification (offline vs online), and the type of data used to train and test the classifiers. We remember that an emotion recognition system has a training phase that should use data that is different from the data used in the test phase.

2.7.1 Classifiers

Since the majority of the works applied more than one classifier, and choose only one for the final configuration of the recognizer, we focus or analysis in the final one. Twenty six different classifiers were selected as the best ones.

In almost 59% of the cases, Support Vector Machines (SVM) was used, with different kernels being applied: Radial Basis Function (RBF) (29.7%), linear (16.2%), polynomial (8.1%), gaussian (5.4%), and pearson (2.7%). Variations such as adaptive SVM, Multiclass Support Vector Machine (ML-SVM) or Least Squares Support Vector Machine (LS-SVM), were used in 8% of these works. Twenty-nine percent of the works that used SVM do not specify the kernel used. The k-Nearest Neighbors (kNN) was selected by almost 14% of the works; some works do not specify the value of k (44.4%), while in the others it varies from k = 2 to 8. Linear Discriminant Analysis (LDA) was used by 6.3% of the authors, while Quadratic Discriminant Analysis (QDA) was selected by 3.2%. Finally, the Naive Bayes (NB) and Multi-Layer Perceptron Back Propagation (MLP-BP) were selected by 6.35% of the authors (3.17% each).

2.7.2 Offline vs online

EEG signals are always changing its nature with time. This non-stationary nature of the signals can lead to classification models, built using specific physiological data, to not reflect the changes that have already occurred to the EEG signals. Most of the classification methods are based on the idea that the data comes from a stationary distribution. Due to this, the classification accuracy is expected to degrade with time unless the model is adapted to reflect the changes occurring in the EEG signals. However, 90% of the works reviewed applied offline classification methods, with only 8% applying online classifica-
tion (more suitable for real-time scenarios). One work applied both online and offline techniques.

2.7.3 User dependence

Another important aspect of the classification process is if the classifier was trained with user-dependent data or not. In the case of user-dependent data, a new model is generated for each user and the testing step is also done with this user data. Typically, better results are obtained, however at the cost of a lack of generalization. In the case of an user-independent model, the data of multiple users are used both for training and testing purposes. This leads to an easier applicability of the model to new users, since there is no need to create a new model. In the works reviewed, 46.8% of them use user-independent data and 43.5% user-dependent data. Around 8% used classifiers trained with models of both types. The rest of the works do not provide any information about their data being user-dependent or user-independent.

2.8 Conclusion

This thesis aims to select informative channels, reliable frequency ranges and effective features for successful EEG emotion recognition. By introducing the statistical modeling based approach along with utilizing neurological ground truths, the feature-channel pairs can be optimized and the number of features and channels can be decreased dramatically leading to a faster and more accurate emotion recognition scheme.
Chapter 3

A Method of Emotion Recognition Based on Statistical Modeling of EMD-Wavelet Transformed EEG Signals Responsive to Music Videos

3.1 Introduction

This Chapter presents a novel emotion recognition scheme which utilizes the selection of proper EEG channels and frequency ranges to recognize emotions in reference with established psychological and neurological ground truths. This ensures the extraction of informative EEG features for robust and accurate classification of the associated emotions. Additionally, in order to effectively retain both time and frequency domain information of non-stationary and non-linear EEG data, an EMD-DWT based method is utilized. Moreover, the best fit statistical models are derived from the empirical mode decomposed and the wavelet transformed EEG signals responsive to music videos, given the fact that a well fitted statistical model can not only help estimate uncertainties in the observed data but also estimate probabilistic future behavior of a system based on past statistical information, thus forming an effective feature set from the selected model parameters.

3.2 Outline of the Proposed Method

In this Chapter, a EMD-wavelet based statistical modeling framework is developed for human emotion recognition using EEG signal. The proposed emotion recognition method consists of three major steps: feature extraction, feature selection and classification. The block diagram of the proposed method is shown in Fig. 3.1.
Fig. 3.1: Block diagram for the proposed method
3.3 Selection of EEG Channels from Neurological Viewpoint

Over the last decade, there has been a significant amount of study about the neurophysiological background of human emotions—how they are evoked and how they become apparent. These studies have shown meaningful correlations between EEG signals and emotions. Motivated by those findings, in this thesis proper EEG channels are selected keeping the neurological background in contention.

According to the studies, there are two main areas of the brain correlated with emotional activity. Those are: the amygdala (located close to the hippocampus, in the frontal portion of the temporal lobe); and the pre-frontal cortex (covers part of the frontal lobe). Although there is no consensus about a possible lateralization of the amygdala, its activation seems to be more related to negative emotions than positive ones [6], [45].

Changes in alpha power and asymmetry between the hemispheres of the brain are related to emotions. Negative emotions can be characterized by a relative right frontal activation, such as fear or disgust. While a relatively greater left frontal activation is associated with positive emotions, such as joy or happiness. This makes the asymmetrical frontal EEG activity a good reflector of changes in the valence dimension [1], [2], [7], [8]. Beta bands are also reported to be related to valence [9]. Pre-frontal asymmetry in the alpha band and temporal asymmetry in gamma band are also present for valence recognition. As for arousal recognition, pre-frontal asymmetry in alpha band and temporal asymmetry in the gamma band are informative [10]. Changes in the gamma band and decrease in alpha wave in different sides of temporal lobe (left for sadness, and right for happiness) are also related with emotions (happiness and sadness). [11], [46]

In summary, we can conclude that the frontal and the frontal portion of the temporal lobes are the most informative about the emotional states, while the alpha, gamma and beta waves appear to be the most discriminating. That is why the EEG channels placed in those areas of the scalp have been selected for feature extraction (see Fig. 4). The selected channels are: Fp1, Fp2, AF3, AF4, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4 and T8. Moreover, eight virtual channels are generated subtracting the signals acquired from left channels from the symmetrical pairs in the right lobe to enhance asymmetrical analysis. This leads to a total of 26 channels to acquire the data to process in the next Sections. The channels are formatted in the following fashion:
Fig. 3.2: Selected channels for the proposed method; Red denotes the channels selected on the left hemisphere, blue denotes the channels selected in the right hemisphere and green denotes the channels selected in the central area of the brain.

All selected channels=[ Selected Channels in left hemisphere, Selected Channels in right hemisphere, Selected Channels in the central area of the brain, Virtual Channels]

### 3.4 EMD-based EEG Signal Decomposition

Empirical mode decomposition (EMD) is a data-driven, adaptive, local signal processing method in which signals are decomposed into narrow band, frequency-modulated oscillations. The EMD operation is a very helpful analysis tool for data that are non-stationary, nonlinear, and too short for standard time-frequency decomposition approaches [5]. In this algorithm, a broadband signal is decomposed into multiple components termed intrinsic mode functions (IMFs) through a data-driven iterative sifting process (Fig. 3.3). The method is computational and performed as follows.
• The upper and lower envelopes of the broadband signal are determined by connecting the maxima and minima respectively, using cubic spline interpolation.

• The mean of the upper and lower envelopes is then subtracted from the original signal. If \( x[n] \) is the target signal and \( m_{1,1}[n] \) is the mean of upper and lower envelopes, then the first component \( h_{1,1}[n] \) can be computed by:

\[
h_{1,1}[n] = x[n] - m_{1,1}[n]
\]  

(1)

• The resulting signal is evaluated as to whether it fulfills the criteria by which an IMF is defined:

a) The number of zero crossings must equal the number of extrema, differing at most by one.

b) The mean of the envelope must be approximately zero.

If the signal does not fulfill these criteria, the procedure, which is referred to as the sifting process, is iterated until an IMF is identified. In each iteration, the output signal from the previous step is considered as the input signal and step I and II are repeated and evaluated by the condition state in step III. If sifting is done for \( k \) iterations, then these steps can be summarized by the following equation:

\[
h_{i,k}[n] = h_{i,k-1}[n] - m_{i,k-1}[n]
\]  

for \( k = 2, 3, 4, \ldots \) Here \( i \) denotes the IMF number and \( k \) denotes the number of iterations. Sifting thus implies repeated removal of lower frequency components, until the remaining signal is an IMF containing the highest frequency oscillations in the data.

• If the output signal after \( k \) iterations satisfies the definition of IMF, \( h_{i,k}[n] \) is stored as the \( i \)-th IMF, \( c_i[n] \) as in

\[
c_i[n] = h_{i,k}[n]
\]  

(3)

The identified IMF is subtracted from the original broadband signal, leaving a residue component, \( r_i[n] \).
Fig. 3.3: An original EEG signal from a subject and the resulting IMFS.

\[ r_i[n] = x[n] - c_i[n] \]  \hspace{1cm} (4)

- the residue signal \( r_i[n] \) is then considered as the target signal \( x[n] \) and the sifting process is repeatedly applied to the residue, resulting in a set of IMFs of decreasing frequency (Fig. 3.3).

- The normalized mean-squared error, \( SD_k \) is calculated between two successive sifting operations and a stopping criterion is defined at which the difference between consecutive IMFs is deemed to be negligible. \( SD_k \) is defined by:

\[ SD_k = \frac{\sum_{n=1}^{N} | h_{k-1}[n] - h_k[n] |^2}{\sum_{n=1}^{N} h_k[n]^2} \]  \hspace{1cm} (5)

### 3.5 Proposed Selection Method of Dominant IMFs

Instead of using all the IMFs, we propose in this thesis to use only the dominant IMFs. The reasons behind this scheme is twofold.
• As discussed in Section 3.2 of this Chapter, the higher frequency components of the EEG data appear to be more discriminating in different emotional states. As a result, using the IMFs containing the highest frequencies of the EEG signal leads to the extraction of effective feature.

• The top three IMFs having the highest frequency are found to contain almost 95% of the total temporal energy of the EEG signal.

As a result, using the dominant high frequency IMFs to extract features not only approximates the EEG signals without losing much of its energy content but also removes the less informative lower frequency components of the EEG signal. IMFs extracted from a EEG signal sorted in descending order according to their temporal energy content and top 3 of them selected as the dominant IMFs. The temporal energy of any IMF is given by:

$$E_d = \sum_{n=1}^{N} c_d[n]^2$$  \hspace{1cm} (6)

Here N and $E_d$ represent the length and temporal of energy of the dominant IMF respectively whereas $c_d[n]$ is the dominant IMF. Fig. 5 shows the IMFs of EEG signal from a channel and only the top 3 are selected for further processing. In order to enhance the discriminating capability and to utilize both the time and frequency domain information content, the selected IMFs are then used for the further analysis via DWT to gain better time- frequency resolution.

3.6 Discrete Wavelet Transform of the Dominant IMFs

Wavelet transforms are a mathematical means for performing signal analysis when signal frequency varies over time. Discrete Wavelet Transform is an implementation of the wavelet transform using a discrete set of wavelet scales and translations [47]. It offers simultaneous localization in time and frequency domain while Fourier transform loses the time information of a signal when converting the signal into frequency domain. The DWT for EEG signal $x[n]$ can be represented in terms of shifted version of a scaling function $\phi$ and a shifted and dilated version of a mother wavelet function $\psi$. The representation of the DWT can be written as:
\[ x[n] = \frac{1}{\sqrt{M}} \sum_{k \in \mathbb{Z}} u_{j_0,k} \phi_{j_0,k}[n] + \sum_{j=-\infty}^{j_0} \sum_{k \in \mathbb{Z}} w_{j,k} \psi_{j,k}[n] \] (7)

where \( M \) is the sample length, \( w_{j,k} \) are the wavelet coefficients and \( u_{j,k} (j < k) \) are the scaling coefficients. Signal decomposition is typically done using the scales \( a = 2, 4, 8, \ldots, 2^L \) with successive approximations being decomposed in turn, so that it is broken down into many lower resolution components, where \( L \) represents the level of decomposition.

In the proposed method, 4 level DWT is performed using db6 wavelet. Coefficients obtained from the wavelet transform of three dominant IMFs are chosen for further analysis since each coefficient represents a particular sub band of EEG signal. Table 3.1 shows the DWT coefficients, their frequency ranges and corresponding sub bands of EEG signal. For further processes, two matrices of DWT coefficients are formed by concatenating side-by-side. One matrix consists of coefficients corresponding to all the frequency bands, while the other one consists of the coefficients describing the alpha, beta and gamma bands only. Both of the matrices are then run through the statistical modeling procedure to look for more distinguishing properties.

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Frequency Range</th>
<th>Sub band</th>
</tr>
</thead>
<tbody>
<tr>
<td>First level detail coefficients (cD1)</td>
<td>32-64 Hz</td>
<td>( \gamma )</td>
</tr>
<tr>
<td>Second level detail coefficients (cD2)</td>
<td>16-32 Hz</td>
<td>( \beta )</td>
</tr>
<tr>
<td>Third level detail coefficients (cD3)</td>
<td>8-16 Hz</td>
<td>( \alpha )</td>
</tr>
<tr>
<td>Fourth level detail coefficients (cD4)</td>
<td>4-8 Hz</td>
<td>( \theta )</td>
</tr>
<tr>
<td>Fourth level approximate coefficients (cA4)</td>
<td>0-4 Hz</td>
<td>( \delta )</td>
</tr>
</tbody>
</table>

### 3.7 Proposed Statistical Modeling of the DWT coefficients

#### 3.7.1 Definition of statistical modeling

A statistical model can be stated as an adequate summary, i.e. a representative smaller version of real world data. There are three broad reasons for using statistical models: causation, prediction/smoothing, and description. A good statistical model should be able to predict/smooth the observed data with an insight into the causal mechanisms in the data and also should work as descriptive summary of its inherent statistical information.
3.7.2 Statistical modeling according to BIC criterion

Although features extracted from the transformed DWT coefficients can efficiently represent the characteristics of the original EEG signal in different details, finding reliable features which represent the EEG data the best, is a difficult task in the field of EEG based emotion recognition. Therefore, we resort to statistical modeling in search of an effective and discriminatory set of features. Both sets of DWT coefficients obtained in the previous step of the procedure are fit to all available distribution functions in MATLAB. After that, the models that describe the data best are found based on Bayesian Information Criterion (BIC). The BIC is formally defined as:

\[
BIC = \ln(n)k - 2\ln(\hat{L})
\]

(8)

where \(x\) = the observed data; \(n\) = the number of data points in \(x\), the number of observations, or equivalently, the sample size; \(k\) = the number of parameters estimated by the model.

When fitting models, addition of parameters may lead to increase the likelihood of fitting the data. But that increases the probability of overfitting the data. BIC attempts to resolve this problem by introducing a penalty term for the number of parameters in the model. The lower the value of BIC, the closer the model is to fitting the dataset perfectly. So, the model with the lowest BIC is chosen as the statistical model that approximates the target data.
In our proposed method, we run the statistical modeling procedure on both sets of DWT coefficients obtained from all sub bands and $\alpha, \beta$ and $\gamma$ subbands of EEG. All the statistical models available in MATLAB are tested and evaluated via BIC criterion. The total number of coefficient vectors in each category (all sub bands and $\alpha, \beta$ and $\gamma$ subbands) is 32 users x 40 videos x 26 channels = 33280. After the statistical modeling process is carried out, it is found that in both cases, almost 95% of all the coefficient vectors are well fitted by only two statistical models. In case of all sub bands' coefficient set, Generalized Extreme Value distribution fits 66.32% of coefficient vectors while 30.04% are fitted by Normal distribution. Whereas, 76.22% of coefficient vectors of $\alpha, \beta$ and $\gamma$ sub bands are well fitted by Extreme Value distribution and 21.45% are best described by Generalized Extreme Value distribution. It is also important to note that the two distributions either came out as the first or the second best performer in terms of fitting the data via BIC. That is why for the sake of creating a feature vector of uniform length The top two statistical models which best describe the transformed coefficient sets are chosen to form the feature vector. The best performing statistical models in both cases are summarized in Fig.3.4. It is evident from the figure that the top two statistical models which best fit the coefficient set of all frequency bands are Generalized extreme value (GEV) distribution and Normal distribution. While for $\alpha, \beta$ and $\gamma$ coefficient set the top two models are Extreme value (EV) distribution and Generalized Extreme Value (GEV) distribution.

While Normal distributions are symmetric, unimodal and asymptotic, extreme value distribution and generalized extreme value distributions deal with the extreme deviations from the median of probability distributions and seeks to assess, from a given ordered sample of a given random variable, the probability of events that are more extreme than any previously observed. Due to the randomness and non-stationary behavior of the EEG data, both normal and extreme value distributions are necessary to successfully and meaningfully describe the sets of DWT coefficients. The statistical models which best describe the coefficient sets thus satisfy all the requirements of statistical modeling by smoothing the data while preserving valuable descriptive information and also by reflecting the non-linear and non-stationery properties of EEG data.
3.8 Proposed Feature Extraction Process

The selected models from the previous Section are selected for feature extraction and the model parameters are used to form the feature sets. The feature sets are constructed as follows:

\[ F_{all} = [k_{GEV}, \sigma_{GEV}, \mu_{GEV}, \mu_{normal}, \sigma_{normal}] \]  
\[ F_{\alpha,\beta,\gamma} = [\mu_{EV}, \sigma_{EV}, k_{GEV}, \sigma_{GEV}, \mu_{GEV}] \]  

For all sub bands’ coefficient set, the number of features per channel per music video is 5. This leads to the formation of a feature set of length 130 from 26 selected channels for one music video. Similarly, for the coefficient set corresponding to \(\alpha, \beta\) and \(\gamma\) frequency bands, the number of features per music video is also 130. Table 3.2 shows the selected models and parameters for feature vector formation. Both of the feature sets are passed on to the next step of the procedure for selection of informative and distinguishing features.

Table 3.2: Statistical models and model parameters for feature vector formation

<table>
<thead>
<tr>
<th>Considered sub bands</th>
<th>Selected Models</th>
<th>Model Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sub bands</td>
<td>GEV</td>
<td>(k) (shape), (\sigma) (scale) and (\mu) (location)</td>
</tr>
<tr>
<td></td>
<td>Normal</td>
<td>(\mu) (mean) and (\sigma) (standard deviation)</td>
</tr>
<tr>
<td>(\alpha, \beta, \gamma) sub bands</td>
<td>EV</td>
<td>(\mu) (location) and (\sigma) (scale)</td>
</tr>
<tr>
<td></td>
<td>GEV</td>
<td>(k) (shape), (\sigma) (scale) and (\mu) (location)</td>
</tr>
</tbody>
</table>

3.9 Effectiveness of the Proposed Features

In this Section, we discuss the effectiveness of the proposed features. To do so, we plot and compare the medians of the feature clusters corresponding to high and low classes across all selected channels and sample videos. As we have two sets of features at our disposal- one corresponding to all sub bands and another to \(\alpha, \beta\) and \(\gamma\) sub bands- we repeat the procedure in both cases. The outcome of this process is demonstrated in the following paragraphs.

3.9.1 All sub-bands’ feature set

As previously stated, In case of all sub bands’ coefficient set we have a total of 130 features per music video which we get by concatenating the five basic features obtained
across all the channels. We first divide the feature set into two groups corresponding to high and low valence classes. Then we further group the features in such a way that each of the five model parameters (our basic features) are regrouped corresponding to the 26 selected channels. In simpler words, we divide the feature set of length 130 into five blocks of length 26 where a single feature from all the channels are grouped together. Now we take each block and calculate the median of the feature vectors one vector at a time. These median values are then compared across high and low valence classes to inspect the class separability and any visible trends in behavior. The same procedure is repeated for high and low arousal, dominance and liking dimensions.

Fig. 3.5-3.9 show the comparison between median of different features in different channels in high and low valence classes. Similarly, Fig. 3.10-3.14 shows the comparison for arousal, Fig. 3.15-3.19 for dominance and 3.20-3.24 for liking. The results of the comparison are summarized in the following table 3.3.

**Fig. 3.5**: Comparison between median of k (shape) parameter of GEV distribution for all sub bands’ coefficient set in high and low valence classes.
Fig. 3.6: Comparison between median of $\sigma$ (scale) parameter of GEV distribution for all sub bands’ coefficient set in high and low valence classes

Fig. 3.7: Comparison between median of $\mu$ (location) parameter of GEV distribution for all sub bands’ coefficient set in high and low valence classes
Fig. 3.8: Comparison between median of $\mu$ (mean) parameter of Normal distribution for all sub bands’ coefficient set in high and low valence classes

Fig. 3.9: Comparison between median of $\sigma$ (standard deviation) parameter of Normal distribution for all sub bands’ coefficient set in high and low valence classes
Fig. 3.10: Comparison between median of $k$ (shape) parameter of GEV distribution for all sub bands’ coefficient set in high and low arousal classes

Fig. 3.11: Comparison between median of $\sigma$ (scale) parameter of GEV distribution for all sub bands’ coefficient set in high and low arousal classes
Fig. 3.12: Comparison between median of $\mu$ (location) parameter of GEV distribution for all sub bands’ coefficient set in high and low arousal classes

Fig. 3.13: Comparison between median of $\mu$ (mean) parameter of Normal distribution for all sub bands’ coefficient set in high and low arousal classes
Fig. 3.14: Comparison between median of $\sigma$ (standard deviation) parameter of Normal distribution for all sub bands’ coefficient set in high and low arousal classes

Fig. 3.15: Comparison between median of $k$ (shape) parameter of GEV distribution for all sub bands’ coefficient set in high and low dominance classes
Fig. 3.16: Comparison between median of $\sigma$ (scale) parameter of GEV distribution for all sub bands’ coefficient set in high and low dominance classes

Fig. 3.17: Comparison between median of $\mu$ (location) parameter of GEV distribution for all sub bands’ coefficient set in high and low dominance classes
Fig. 3.18: Comparison between median of $\mu$ (mean) parameter of Normal distribution for all sub bands’ coefficient set in high and low dominance classes

Fig. 3.19: Comparison between median of $\sigma$ (standard deviation) parameter of Normal distribution for all sub bands’ coefficient set in high and low dominance classes
Fig. 3.20: Comparison between median of k (shape) parameter of GEV distribution for all sub bands’ coefficient set in high and low liking classes

Fig. 3.21: Comparison between median of \( \sigma \) (scale) parameter of GEV distribution for all sub bands’ coefficient set in high and low liking classes
Fig. 3.22: Comparison between median of $\mu$ (location) parameter of GEV distribution for all sub bands’ coefficient set in high and low liking classes

Fig. 3.23: Comparison between median of $\mu$ (mean) parameter of Normal distribution for all sub bands’ coefficient set in high and low liking classes
Fig. 3.24: Comparison between median of $\sigma$ (standard deviation) parameter of Normal distribution for all sub bands’ coefficient set in high and low liking classes

### 3.9.2 $\alpha$, $\beta$ and $\gamma$ sub bands’ feature set

In this Subsection, the same comparisons are made for the $\alpha$, $\beta$ and $\gamma$ sub bands’ feature set as was done for all sub bands’ feature set. The outcomes are summarized in Fig. 3.25-3.44. Fig. 3.25-3.29 show the comparison between median of different features in different channels in high and low valence classes. Similarly, Fig. 3.30-3.34 shows the comparison for arousal, Fig. 3.35-3.39 for dominance and 3.40-3.44 for liking. Same as before, the changes in median of features in different emotion dimensions across different channels are summarized in the table 3.4 afterwards.
<table>
<thead>
<tr>
<th>Emotion Dimension</th>
<th>Feature Name</th>
<th>Channels showing decrease in feature value from high to low class</th>
<th>Channels showing increase in feature value from high to low class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>$k_{GEV}$</td>
<td>Fp1, AF3, F3, T7, Fp2, F8, FC2, F3-F4, FC5-FC6, T7-T8, C3-C4</td>
<td>FC5, AF4, F4, FC6, T8, Cz</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>AF3, T7, C3, F4, F8, FC2, FC6, T8, C4</td>
<td>Fp1, F3, FC5, Fp1-Fp2, AF3-AF4</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>AF4, C3-C4</td>
<td>F3, F7, T7, C3, FC2, Fz, F3-F4, F7-F8</td>
</tr>
<tr>
<td></td>
<td>$\mu_{Normal}$</td>
<td>Fp1, AF3, F3, FC1, FC5, T7, AF4, FC1-FC2, T7-T8, C3-C4</td>
<td>F7, FC2, C4, Fz, Cz</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Normal}$</td>
<td>Fp1, AF3, F3, FC5, T7, C3, Fp2, T8, Fz, Fz, Cz, AF3- AF4, F3-F4, F7-T8</td>
<td>FC1, AF4, F8, FC2, Fp1-Fp2, F7-F8</td>
</tr>
<tr>
<td>Arousal</td>
<td>$k_{GEV}$</td>
<td>T7, AF4, F4, F8, FC2, FC6, C4, Fz</td>
<td>F3, Fp2, T8, Cz, Fp1-Fp2, F7-F8</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>F3, F7, AF4, C4, Fz</td>
<td>Fp1, AF3, FC1, Fp2, F4, F8, FC6, T8</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>AF3, F7, AF4, Fp1-Fp2, AF3-AF4, FC5-FC6, C3-C4</td>
<td>Fp1, FC1, Fp2, F4, FC2, T8, C4, Fz, Cz</td>
</tr>
<tr>
<td></td>
<td>$\mu_{Normal}$</td>
<td>AF3, F7, FC1, T7, AF4, FC5, FC6, T8, Fp1-Fp2, FC1-FC2, T7-T8</td>
<td>Pp1, C3, Fz, Cz, AF3-AF4, FC5-FC6, C3-C4</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Normal}$</td>
<td>Fp1, F3, FC1, FC5, T7, C3, Fp2, F4, F8, FC2, FC6, T8, C4, Fz, Cz, F7-F8</td>
<td>AF4, Fp1-Fp2, AF3-AF4, FC5-FC6, T7-T8</td>
</tr>
<tr>
<td>Dominance</td>
<td>$k_{GEV}$</td>
<td>AF3, FC5, T7, Fp2, AF4, F8, FC2, T8, C4, Fz</td>
<td>F3, Fp2, T8, Cz, Fp1-Fp2, F7-F8</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>AF4, F4, F8, FC2, FC6, C4, F7-F8</td>
<td>F3, Fp2, T8, Cz, Fp1-Fp2, F7-F8</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>AF3, FC1, AF4, Fz, F7-F8, FC5-FC6</td>
<td>F3, T7, C3, F4, FC2, FC6, T8, Cz, F3-F4, C3-C4</td>
</tr>
<tr>
<td></td>
<td>$\mu_{Normal}$</td>
<td>Fp1, AF3, F3, FC1, FC5, T7, AF4, FC6, T8, Fz, FC1-FC2, T7-T8, C3-C4</td>
<td>F7, C3, Fp2, C4, Cz, AF3-AF4</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Normal}$</td>
<td>F3, T7, C3, Fp2, FC2, FC6, T8</td>
<td>AF3, FC1, F8, Fp1-Fp2, F7-F8</td>
</tr>
<tr>
<td>Liking</td>
<td>$k_{GEV}$</td>
<td>F7, FC5, Fp2, Fz, F3-F4, FC5-FC6, T7-T8</td>
<td>FC1, T7, C3, AF4, F4, FC2, FC6, C4, C2</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>AF3, FC1, T7, AF4, F4, F8, FC2, FC6</td>
<td>F3, F7, AF3-AF4, F3-F4, F7-F8, FC5-FC6</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>FC5-FC6, C3-C4</td>
<td>Fp1, F3, F7, FC1, FC5, T7, C3, Fp2, F4, F8, FC6, C4, Fz, Cz, F3-F4</td>
</tr>
<tr>
<td></td>
<td>$\mu_{Normal}$</td>
<td>Fp1, AF3, F7, Fp2, F4, Cz, FC1-FC2</td>
<td>FC1, FC5, C3, FC2, T8, C4, AF3-AF4, T7-T8, C3-C4</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Normal}$</td>
<td>Fp1, AF3, F3, F7, FC1, FC5, C3, Fp2, AF4, F4, F8, FC2, FC6, C4, Fz, Cz, Fp1-Fp2, AF3-AF4, F3-F4, F7-F8</td>
<td>---</td>
</tr>
</tbody>
</table>
Fig. 3.25: Comparison between median of \( \mu \) (location) parameter of EV distribution for \( \alpha, \beta \) and \( \gamma \) sub bands’ coefficient set in high and low valence classes

Fig. 3.26: Comparison between median of \( \sigma \) (scale) parameter of EV distribution for \( \alpha, \beta \) and \( \gamma \) sub bands’ coefficient set in high and low valence classes
Fig. 3.27: Comparison between median of $k$ (shape) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands’ coefficient set in high and low valence classes

Fig. 3.28: Comparison between median of $\sigma$ (scale) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands’ coefficient set in high and low valence classes
Fig. 3.29: Comparison between median of $\mu$ (location) parameter of GEV distribution for $\alpha$, $\beta$ and $\gamma$ sub bands’ coefficient set in high and low valence classes

Fig. 3.30: Comparison between median of $\mu$ (location) parameter of EV distribution for $\alpha$, $\beta$ and $\gamma$ sub bands’ coefficient set in high and low arousal classes
Fig. 3.31: Comparison between median of $\sigma$ (scale) parameter of EV distribution for $\alpha, \beta$ and $\gamma$ sub bands’ coefficient set in high and low arousal classes

Fig. 3.32: Comparison between median of $k$ (shape) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands’ coefficient set in high and low arousal classes
Fig. 3.33: Comparison between median of $\sigma$ (scale) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands’ coefficient set in high and low arousal classes

Fig. 3.34: Comparison between median of $\mu$ (location) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands’ coefficient set in high and low arousal classes
Fig. 3.35: Comparison between median of $\mu$ (location) parameter of EV distribution for $\alpha$, $\beta$ and $\gamma$ sub bands' coefficient set in high and low dominance classes.

Fig. 3.36: Comparison between median of $\sigma$ (scale) parameter of EV distribution for $\alpha$, $\beta$ and $\gamma$ sub bands' coefficient set in high and low dominance classes.
Fig. 3.37: Comparison between median of \( k \) (shape) parameter of GEV distribution for \( \alpha, \beta \) and \( \gamma \) sub bands’ coefficient set in high and low dominance classes

Fig. 3.38: Comparison between median of \( \sigma \) (scale) parameter of GEV distribution for \( \alpha, \beta \) and \( \gamma \) sub bands’ coefficient set in high and low dominance classes
Fig. 3.39: Comparison between median of $\mu$ (location) parameter of GEV distribution in high and low dominance classes

Fig. 3.40: Comparison between median of $\mu$ (location) parameter of EV distribution for $\alpha$, $\beta$ and $\gamma$ sub bands’ coefficient set in high and low liking classes
Fig. 3.41: Comparison between median of $\sigma$ (scale) parameter of EV distribution for $\alpha, \beta$ and $\gamma$ sub bands’ coefficient set in high and low liking classes

Fig. 3.42: Comparison between median of $k$ (shape) parameter of GEV distribution for $\alpha, \beta$ and $\gamma$ sub bands’ coefficient set in high and low liking classes
Fig. 3.43: Comparison between median of $\sigma$ (scale) parameter of GEV distribution for $\alpha$, $\beta$ and $\gamma$ sub bands’ coefficient set in high and low liking classes

Fig. 3.44: Comparison between median of $\mu$ (location) parameter of GEV distribution for $\alpha$, $\beta$ and $\gamma$ sub bands’ coefficient set in high and low liking classes
Table 3.4: Comparison of median of feature values between high and low classes for α, β and γ sub bands’ feature set

<table>
<thead>
<tr>
<th>Emotion Dimension</th>
<th>Feature Name</th>
<th>Channels showing decrease in feature value from high to low class</th>
<th>Channels showing increase in feature value from high to low class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>$\mu_{EV}$</td>
<td>Fp1,F3,F7,FC6,T8,C4,Fz,Cz,Fp1-Fp2,AF3-AF4,F3-F4,F7-F8,FC1-FC2,T7-T8,C3-C4</td>
<td>Fp2,F8,FC2</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{EV}$</td>
<td>F3,FC5,C3,T8,C4,Fz,AF3-AF4,F3-F4,F7-F8,FC1-FC2,C3-C4</td>
<td>AF3,FC1,T7,Fp2,F8,FC2,Cz</td>
</tr>
<tr>
<td></td>
<td>$k_{GEV}$</td>
<td>Fp2,F8,FC6,T7-T8</td>
<td>F3,F7,T7,AF4,F4,FC2,T8,Fz,Cz</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>Fp1,F3,FC5,AF4,T8,Cz,AF3-AF4,F3-F4,F7-F8,FC5-FC6,C3-C4</td>
<td>F7,FC1,T7,Fp2,F4,F8,FC2</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>Fp2,F4,F8,FC2</td>
<td>F3,FC6,T8,C4,Fz,Cz,Fp1-Fp2,AF3-AF4,F3-F4,F7-F8,FC1-FC2,T7-T8,C3-C4</td>
</tr>
<tr>
<td>Arousal</td>
<td>$\mu_{EV}$</td>
<td>AF3,FC1,T7,C3,F8,T8,C4,Fz,AF3-AF4</td>
<td>F7,AF4,FC2,Cz,F3-F4,FC1-FC2,FC5-FC6,C3-C4</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{EV}$</td>
<td>Fp1,AF3,FC1,FC5,T7,F8,T8,C4,Fz</td>
<td>C3,AF4,FC2,Cz,AF3-AF4,F3-F4,FC1-FC2</td>
</tr>
<tr>
<td></td>
<td>$k_{GEV}$</td>
<td>AF3,F7,AF4,Cz,F3-F4,F7-F8,T7-T8,C3-C4</td>
<td>F3,FC1,T7,Fp2,F4,FC6,C4,Fz</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>Fp1,AF3,FC1,FC5,F4,F8,T8,C4,Fz</td>
<td>F7,AF4,FC2,Cz,F3-F4,F7-F8,FC5-FC6,C3-C4</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>F7,AF4,FC2,Cz,F3-F4,FC1-FC2,FC5-FC6,C3-C4</td>
<td>Fp1,AF3,FC5,T7,T8,C4,Fz,AF3-AF4,T7-T8</td>
</tr>
<tr>
<td>Dominance</td>
<td>$\mu_{EV}$</td>
<td>F3,T7,Fp2,AF4,FC6,T8,Fz,Cz,FC1-FC2,C3-C4</td>
<td>F4,FC2,FC4,AF3-AF4,F7-F8</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{EV}$</td>
<td>Fp1,F3,FC1,T7,C3,FC6,Fz,FC1-FC2</td>
<td>AF3,AF3-AF4</td>
</tr>
<tr>
<td></td>
<td>$k_{GEV}$</td>
<td>AF3,Fp2,F8,FC2,AF3-AF4,F7-F8,T7-T8</td>
<td>F3,F7,FC1,FC5,T7,C4,Fz,Cz,Fp1-Fp2</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>F3,T7,Fp2,AF4,FC6,Cz,C3-C4</td>
<td>F7,F4,FC2,C4,AF3-AF4,F7-F8,FC5-FC5</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>F4,FC2,AF3-AF4,F7-F8</td>
<td>F3,FC5,T7,Fp2,FC6,T8,Fz,Cz,FC1-FC2,T7-T8,C3-C4</td>
</tr>
<tr>
<td>Liking</td>
<td>$\mu_{EV}$</td>
<td>Fp1,F7,C3,F4,FC6,T8,C4,Fz,Cz,Fp1-Fp2,AF3-AF4,F7-F8,C3-C4</td>
<td>T7,Fp2,AF4</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{EV}$</td>
<td>Fp1,FC1,FC5,T8,C4,Fz,AF3-AF4,F7-F8</td>
<td>T7,FC2,Cz,FC1-FC2</td>
</tr>
<tr>
<td></td>
<td>$k_{GEV}$</td>
<td>AF3,Fp2,FC6,T8,FC1-FC2,T7-T8</td>
<td>Fp1,F3,F7,AF4,FC2,C4,Fz,Cz,F3-F4</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>Fp1,F8,FC2,FC6,C4,Fz,AF3-AF4,F3-F4,FC5-FC6,C3-C4</td>
<td>T7,Fp2,AF4,T8</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>T7,Fp2,AF4</td>
<td>F7,FC1,F4,FC6,T8,C4,Fz,Cz,Fp1-Fp2,AF3-AF4,F7-F8,T7-T8</td>
</tr>
</tbody>
</table>
3.9.3 Euclidean distance between median values of features of different classes

In the previous two Subsections, an overall idea about the inter class separability is rendered to the reader. However, to obtain a numeric metric about the class separability of various features we calculate the euclidean distance between median values of features of different classes. The values of distance are normalized by dividing by the range of features values for comparability. The result of the comparison is stated in the following table 3.5.

Table 3.5: Comparison of Euclidean distance between median of feature values of high and low classes for all sub bands’ and $\alpha, \beta$ and $\gamma$ sub bands’ feature sets expressed as percentage of range of feature values)

<table>
<thead>
<tr>
<th>Emotion Dimension</th>
<th>Feature Name</th>
<th>All sub bands(%)</th>
<th>Feature Name</th>
<th>$\alpha, \beta$ and $\gamma$ sub bands'(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Valence</strong></td>
<td>$k_{GEV}$</td>
<td>7.91</td>
<td>$\mu_{EV}$</td>
<td>35.59</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>10.03</td>
<td>$\sigma_{EV}$</td>
<td>15.49</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>9.51</td>
<td>$k_{GEV}$</td>
<td>9.27</td>
</tr>
<tr>
<td></td>
<td>$\mu_{Normal}$</td>
<td>11.13</td>
<td>$\sigma_{GEV}$</td>
<td>22.22</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Normal}$</td>
<td>9.06</td>
<td>$\mu_{GEV}$</td>
<td>35.41</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>6.90</td>
<td>Overall</td>
<td>21.17</td>
</tr>
<tr>
<td><strong>Arousal</strong></td>
<td>$k_{GEV}$</td>
<td>10.41</td>
<td>$\mu_{EV}$</td>
<td>31.16</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>5.75</td>
<td>$\sigma_{EV}$</td>
<td>12.91</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>10.26</td>
<td>$k_{GEV}$</td>
<td>13.15</td>
</tr>
<tr>
<td></td>
<td>$\mu_{Normal}$</td>
<td>13.91</td>
<td>$\sigma_{GEV}$</td>
<td>19.67</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Normal}$</td>
<td>10.98</td>
<td>$\mu_{GEV}$</td>
<td>32.08</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>6.32</td>
<td>Overall</td>
<td>19.01</td>
</tr>
<tr>
<td><strong>Dominance</strong></td>
<td>$k_{GEV}$</td>
<td>10.42</td>
<td>$\mu_{EV}$</td>
<td>32.13</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>14.43</td>
<td>$\sigma_{EV}$</td>
<td>15.01</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>10.13</td>
<td>$k_{GEV}$</td>
<td>15.96</td>
</tr>
<tr>
<td></td>
<td>$\mu_{Normal}$</td>
<td>16.46</td>
<td>$\sigma_{GEV}$</td>
<td>18.77</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Normal}$</td>
<td>11.46</td>
<td>$\mu_{GEV}$</td>
<td>32.33</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>7.93</td>
<td>Overall</td>
<td>19.75</td>
</tr>
<tr>
<td><strong>Liking</strong></td>
<td>$k_{GEV}$</td>
<td>6.97</td>
<td>$\mu_{EV}$</td>
<td>37.25</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>9.95</td>
<td>$\sigma_{EV}$</td>
<td>14.70</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>8.32</td>
<td>$k_{GEV}$</td>
<td>8.83</td>
</tr>
<tr>
<td></td>
<td>$\mu_{Normal}$</td>
<td>14.50</td>
<td>$\sigma_{GEV}$</td>
<td>27.87</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Normal}$</td>
<td>8.87</td>
<td>$\mu_{GEV}$</td>
<td>34.09</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>7.25</td>
<td>Overall</td>
<td>22.19</td>
</tr>
</tbody>
</table>

The values in the table indicate that over all the dimensions of emotion, $\alpha, \beta$ and $\gamma$ sub bands’ feature set show greater inter class separability in terms of distance between median of feature values of high and low classes because the distance is greater than that
of all sub bands’ feature set across all emotion dimensions.

### 3.9.4 Measure of intra class compactness

In previous Subsection, measure of inter class separability was demonstrated in terms of distance between median of feature values of high and low classes. In this part of the Chapter, we will calculate and compare intra class compactness of high and low classes of all emotion dimensions for both all sub bands’ and \( \alpha, \beta \) and \( \gamma \) sub bands’ feature sets. For this purpose, we calculated the euclidean distance between all the feature values and the median of the feature values. After that, we averaged that distance over all the channels. The obtained values are then expressed as percentage of range of feature values and used as a metric of intra class compactness. The lower the value of this metric, the more compact the features are in a particular class. The outcome of this process is summarized in the following table 3.6 and table 3.7.

Table 3.6: Comparison of Euclidean distance between all feature values and median of feature values of high and low classes for all sub bands’ feature set expressed as percentage of range of feature values

<table>
<thead>
<tr>
<th>Emotion Dimension</th>
<th>Feature Name</th>
<th>Compactness of high class</th>
<th>Compactness of low class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>( k_{GEV} )</td>
<td>16.13</td>
<td>15.68</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{GEV} )</td>
<td>17.13</td>
<td>17.45</td>
</tr>
<tr>
<td></td>
<td>( \mu_{GEV} )</td>
<td>18.83</td>
<td>18.79</td>
</tr>
<tr>
<td></td>
<td>( \mu_{Normal} )</td>
<td>16.00</td>
<td>14.99</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{Normal} )</td>
<td>19.23</td>
<td>19.12</td>
</tr>
<tr>
<td>Arousal</td>
<td>( k_{GEV} )</td>
<td>15.78</td>
<td>16.31</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{GEV} )</td>
<td>17.21</td>
<td>17.36</td>
</tr>
<tr>
<td></td>
<td>( \mu_{GEV} )</td>
<td>18.78</td>
<td>18.86</td>
</tr>
<tr>
<td></td>
<td>( \mu_{Normal} )</td>
<td>15.46</td>
<td>15.92</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{Normal} )</td>
<td>19.14</td>
<td>19.32</td>
</tr>
<tr>
<td>Dominance</td>
<td>( k_{GEV} )</td>
<td>16.26</td>
<td>15.27</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{GEV} )</td>
<td>17.16</td>
<td>17.48</td>
</tr>
<tr>
<td></td>
<td>( \mu_{GEV} )</td>
<td>19</td>
<td>18.45</td>
</tr>
<tr>
<td></td>
<td>( \mu_{Normal} )</td>
<td>15.73</td>
<td>15.49</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{Normal} )</td>
<td>19.32</td>
<td>18.91</td>
</tr>
<tr>
<td>Liking</td>
<td>( k_{GEV} )</td>
<td>15.95</td>
<td>16.04</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{GEV} )</td>
<td>17.24</td>
<td>17.32</td>
</tr>
<tr>
<td></td>
<td>( \mu_{GEV} )</td>
<td>18.87</td>
<td>18.70</td>
</tr>
<tr>
<td></td>
<td>( \mu_{Normal} )</td>
<td>15.89</td>
<td>15.05</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{Normal} )</td>
<td>19.24</td>
<td>19.09</td>
</tr>
</tbody>
</table>
Table 3.7: Comparison of Euclidean distance between all feature values and median of feature values of high and low classes for $\alpha$, $\beta$ and $\gamma$ sub bands’ feature set expressed as percentage of range of feature values

<table>
<thead>
<tr>
<th>Emotion Dimension</th>
<th>Feature Name</th>
<th>Compactness of high class</th>
<th>Compactness of low class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>$\mu_{EV}$</td>
<td>25.18</td>
<td>24.20</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{EV}$</td>
<td>21.09</td>
<td>20.50</td>
</tr>
<tr>
<td></td>
<td>$k_{GEV}$</td>
<td>18.36</td>
<td>18.65</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>22.77</td>
<td>21.83</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>25.28</td>
<td>24.39</td>
</tr>
<tr>
<td>Arousal</td>
<td>$\mu_{EV}$</td>
<td>24.77</td>
<td>24.92</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{EV}$</td>
<td>20.73</td>
<td>21.14</td>
</tr>
<tr>
<td></td>
<td>$k_{GEV}$</td>
<td>18.70</td>
<td>18.20</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>22.26</td>
<td>22.74</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>24.9</td>
<td>20.04</td>
</tr>
<tr>
<td>Dominance</td>
<td>$\mu_{EV}$</td>
<td>24.79</td>
<td>24.80</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{EV}$</td>
<td>20.95</td>
<td>20.76</td>
</tr>
<tr>
<td></td>
<td>$k_{GEV}$</td>
<td>18.70</td>
<td>18.08</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>22.57</td>
<td>22.06</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>24.88</td>
<td>24.95</td>
</tr>
<tr>
<td>Liking</td>
<td>$\mu_{EV}$</td>
<td>24.91</td>
<td>24.69</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{EV}$</td>
<td>21.09</td>
<td>20.43</td>
</tr>
<tr>
<td></td>
<td>$k_{GEV}$</td>
<td>18.53</td>
<td>18.59</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{GEV}$</td>
<td>22.62</td>
<td>22.03</td>
</tr>
<tr>
<td></td>
<td>$\mu_{GEV}$</td>
<td>25.01</td>
<td>24.81</td>
</tr>
</tbody>
</table>

The value of normalized euclidean distances presented over the previous two tables suggest that all sub bands’ feature set is lightly more compact than the $\alpha$, $\beta$ and $\gamma$ sub bands’ feature set. To further enhance the distinguishing ability of the features, we should apply some feature selection scheme in order to get rid of redundant features to achieve better inter class separability and intra class compactness. The following Section discusses the feature selection scheme applied in the proposed method.

### 3.10 Proposed Feature Selection Method

In this work we have used the ReliefF algorithm for feature selection [48]. ReliefF is an extension of previously developed Relief algorithm [49]. Relief is not dependent on heuristics, runs in low-order polynomial time, and is noise-tolerant and robust to feature interactions. It is simple and has low computational time. However, it does not behave well with small set of training instances. [50] To address this, ReliefF runs the outer loop of the algorithm over all available training instances [48]. In addition, it can be extended
to the multi-class problem.

ReliefF calculates a feature score for each feature which can then be applied to rank and select top scoring features for feature selection. The algorithm penalizes the predictors that give different values to neighbors of the same class, and rewards predictors that give different values to neighbors of different classes. So the use of ReliefF function ensures better class separability and results in better classification performance. The feature selection procedure is carried out as follows:

- Let \( x \) be the dimension of input information feature set \( A \) as obtained in Section 3.8 \((m \geq l)\), where \( m \) and \( l \) represent no of rows and columns respectively.

- The objective is to find an \( m \times g \) subset \( A_1 \) \((g \leq l)\) of \( A \), which contains significant part of the information of feature matrix \( A \).

- MATLAB function ReliefF is first used to determine the ranks of the features and their corresponding scores.

- All the predictors having negative scores are excluded from the feature set due to their poor performance in predicting the correct label assigned to a particular feature set.

- Among the features having positive scores, the features which have very low positive score have almost equal success and failure rate. Using those features may lead to erroneous classifier performance. To counter this, \( g \) number of features for which the scores of features exhibit 99% of the total score are obtained and selected to form the final feature set.

- The reduced model discards \((l - g)\) number of low ranked features based on the outcome of ReliefF scheme.

3.11 Change in Inter class Separability and Intra Class Compactness

In this Section, we briefly discuss the inter class separability and intra class compactness after feature selection is applied. For comparison, we use the same metrics as the previous Section. The results are summarized in the following tables 3.9 and 3.10.
Table 3.8: Change in inter class separability and intra class compactness for all sub bands’ feature set after feature selection

<table>
<thead>
<tr>
<th>Emotion Dimension</th>
<th>Before/after feature selection</th>
<th>High class compactness</th>
<th>Low class compactness</th>
<th>Inter class separability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>Before feature selection</td>
<td>17.466</td>
<td>17.207</td>
<td>6.90</td>
</tr>
<tr>
<td></td>
<td>After feature selection</td>
<td>19.2</td>
<td>19.26</td>
<td>21.44</td>
</tr>
<tr>
<td>Arousal</td>
<td>Before feature selection</td>
<td>17.27</td>
<td>17.55</td>
<td>6.32</td>
</tr>
<tr>
<td></td>
<td>After feature selection</td>
<td>18.70</td>
<td>19.21</td>
<td>21.66</td>
</tr>
<tr>
<td>Dominance</td>
<td>Before feature selection</td>
<td>17.49</td>
<td>17.12</td>
<td>7.93</td>
</tr>
<tr>
<td></td>
<td>After feature selection</td>
<td>20.68</td>
<td>20.70</td>
<td>21.00</td>
</tr>
<tr>
<td>Liking</td>
<td>Before feature selection</td>
<td>17.44</td>
<td>17.24</td>
<td>7.25</td>
</tr>
<tr>
<td></td>
<td>After feature selection</td>
<td>19.06</td>
<td>18.87</td>
<td>20.84</td>
</tr>
</tbody>
</table>

Table 3.9: Change in inter class separability and intra class compactness for α, β and γ sub bands’ feature set after feature selection

<table>
<thead>
<tr>
<th>Emotion Dimension</th>
<th>Before/after feature selection</th>
<th>High class compactness</th>
<th>Low class compactness</th>
<th>Inter class separability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>Before feature selection</td>
<td>22.59</td>
<td>21.86</td>
<td>21.17</td>
</tr>
<tr>
<td></td>
<td>After feature selection</td>
<td>24.10</td>
<td>23.42</td>
<td>54.27</td>
</tr>
<tr>
<td>Arousal</td>
<td>Before feature selection</td>
<td>22.27</td>
<td>22.41</td>
<td>19.01</td>
</tr>
<tr>
<td></td>
<td>After feature selection</td>
<td>23.54</td>
<td>23.84</td>
<td>50.54</td>
</tr>
<tr>
<td>Dominance</td>
<td>Before feature selection</td>
<td>22.38</td>
<td>22.13</td>
<td>19.75</td>
</tr>
<tr>
<td></td>
<td>After feature selection</td>
<td>25.79</td>
<td>25.11</td>
<td>53.95</td>
</tr>
<tr>
<td>Liking</td>
<td>Before feature selection</td>
<td>22.43</td>
<td>22.11</td>
<td>22.19</td>
</tr>
<tr>
<td></td>
<td>After feature selection</td>
<td>24.1</td>
<td>23.77</td>
<td>58.49</td>
</tr>
</tbody>
</table>

The results shown in the previous tables clearly shows that the feature selection method greatly enhanced the inter class separability for both all sub bands’ and α, β and γ sub bands’ feature sets with a slight decline in intra class compactness. It is also evident that the α, β and γ sub bands’ feature set has way greater inter class distance which should lead to better classification accuracies.

3.12 Discussion on Selected Features

The total number of features per music video is 130 for both all sub bands’ and α, β and γ sub bands’ coefficient sets. By using ReliefF, redundant features are removed and the resulting number of features are reported in table 3.8. The weights assigned to different features are shown in the bar charts presented in Fig. 3.45 and Fig. 3.46. Fig. 3.45 demonstrates the feature weights in different emotion dimensions obtained from all sub
bands’ coefficient set. And Fig. 3.46 shows the feature weights obtained from utilizing only $\alpha, \beta$ and $\gamma$ sub bands’ coefficients. From the table and the figures, we see that less number of features obtained from all frequency sub bands have significant positive score assigned to them than the features obtained from $\alpha, \beta$ and $\gamma$ sub bands. This is an indication to the conclusion that $\alpha, \beta$ and $\gamma$ sub bands are more discriminating and holds more information about human emotion content. We hypothesize that using the features obtained from $\alpha, \beta$ and $\gamma$ sub bands will lead to a better classification performance in different emotion dimensions namely valence, arousal, dominance and liking. We verify this hypothesis in the simulation results and performance comparison Section of this thesis.

Table 3.10: Number of features selected using ReliefF in different emotion dimension

<table>
<thead>
<tr>
<th>Emotion dimension</th>
<th>Features selected for all sub bands</th>
<th>Features selected for $\alpha, \beta$ and $\gamma$ sub bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>24</td>
<td>48</td>
</tr>
<tr>
<td>Arousal</td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td>Dominance</td>
<td>31</td>
<td>41</td>
</tr>
<tr>
<td>Liking</td>
<td>81</td>
<td>105</td>
</tr>
</tbody>
</table>

Fig. 3.45: Weights of features obtained from DWT coefficients corresponding to all sub bands
3.13 Classification Methods

In this thesis, we explore the performance of different classifiers such as: Support Vector Machine (SVM), K-nearest neighbors (KNN), Naive Bayes (NB), Decision tree and Ensemble.

3.13.1 Support Vector Machine (SVM)

SVM [51] is a binary classifier which can be extended into a multi-class classifier as well. Its ability to robustly generalize data makes it capable to work well in different classification problems. That is why SVM has gained popularity in the research arena.

Consider a training set \((x_j, y_j), 1 \leq j \leq N\), where \(x_j\) denotes the feature vectors extracted from EEG signals, \(y_j\) denotes the corresponding emotion labels, and \(N\) is the number of data.

The SVM decision function can be written as

\[
f(x) = \sum_{n=1}^{N} \alpha_n y_n k(s_n, x) + b
\]  

(11)

where \(x\) is the input vector (the feature vector extracted from EEG signals), \(k\) is the kernel function, \(s_n\) denotes support vectors, \(\alpha_n\) are the weights and \(b\) is the bias. Computing the SVM classifier amounts to minimizing an expression of the form:
\[
\frac{1}{n} \sum_{i=1}^{n} \max(0, 1 - y_i(w \cdot x_i - b)) + \lambda \|w\|^2
\] (12)

where \( w \) is the normal vector to the hyperplane acting as a decision boundary between the binary classes. In the scope of this thesis, we used the Radial Basis Function (RBF) as kernel. To train the SVM, weights \( \alpha \) are found for existing data such that

\[
f(x) = \begin{cases} 
\geq 0 & y_i = 1 \\
< 0 & y_i = 0
\end{cases}
\] (13)

where 1 and 0 denote positive and negative emotion classes, respectively.

### 3.13.2 K-nearest Neighbors (KNN)

KNN is another classifier which has been popularly implemented in emotion recognition problems [52]. The classification is based on user-defined constant integer \( k \), where new case will be assigned to the class most common amongst its \( k \) nearest neighbors measured by a distance metric. However, the KNN classification is prone to predicting classes with more examples dominating the classification while training imbalanced datasets. For the purpose of this thesis, MATLAB inbuilt function fitcknn is used to fit KNN model to the data. This function attempts to minimize the cross-validation loss for the fitcknn by varying its parameters, including the number of neighbors and distance metric depending on the dataset. The distance metrics available to this function are City block distance, Chebychev distance, Minkowski distance, Euclidean and standardised Euclidean distance, Hamming distance, Jaccard coefficient, Mahalanobis distance, and Spearman’s rank correlation.

### 3.13.3 Naive Bayes (NB)

NB classifiers [53], [54] are a family of simple probabilistic classifiers that assume strong independence among all input variables. The aim of NB classifier is to calculate the conditional probability that data points belong to a specific class given the input features and chooses the class with the highest probability. Thus, the goal of an NB classifier is to find the probability \( p(C \mid F_1, ..., F_n) \), where \( C \) is the class variable and \( F_1, ..., F_n \) are the data points. As this is a difficult probability to compute, that is why the classifier resorts to the Bayes theorem. Furthermore, all input variables \( F_i \) are assumed to be independent.
Hence, the conditional probability that a given data point belong to a specific class given the input features can be written as

\[ p(C \mid F_1, \ldots, F_n) = \frac{p(C) \prod_{i=1}^{n} p(F_i \mid C)}{p(F_1, \ldots, F_n)} \]  

(5)

This can be written as:

\[ p(C \mid F_1, \ldots, F_n) = \frac{1}{Z} p(C) \prod_{i=1}^{n} p(F_i \mid C) \]  

(6)

where \( Z \) is a scaling factor dependent on \( F_1, \ldots, F_n \). Despite the assumption of strong independence among features, an NB classifier still performs surprisingly well even when the assumption does not hold adequately true.

### 3.13.4 Decision Tree

Decision Tree Classifier is a simple and widely used classification technique. It poses a series of carefully crafted questions about the attributes of the test record. Each time it receives an answer, a follow-up question is asked until a conclusion about the class label of the record is reached. In each step of this procedure all the features are considered and different split points are tried and tested using a cost function and the split with the lowest cost is selected. In classification, a commonly used cost function is the Gini score which is given by:

\[ G = \sum_k p_k (1 - p_k) \]  

(7)

Here, \( p_k \) is proportion of same class inputs present in a particular group. However, while working on a classification problem which has a large set of features, decision tree classifier results in large number of splits, which in turn gives a huge tree. Such trees are complex and can easily lead to over fitting. It can be prevented by setting a minimum number of training inputs to use on each leaf or by setting maximum depth of the model. The performance of a tree can be further increased by pruning which involves removing the branches that make use of features having low importance.

### 3.13.5 Ensemble Classifier

Ensemble methods are meta-algorithms that combine several machine learning techniques into one predictive model in order to decrease variance (bagging), bias (boosting), or im-
prove predictions (stacking). This approach usually allows the production of better predictive performance compared to a single model. Ensemble methods can be divided into two groups: sequential ensemble methods (e.g. AdaBoost) and parallel ensemble methods (e.g. Random Forest). In order for ensemble methods to be more accurate than any of its individual members, the base learners have to be as accurate as possible and as diverse as possible. Bagging, boosting and stacking plays an important role in that process. Bagging uses bootstrap sampling to obtain the data subsets for training the base learners. For aggregating the outputs of base learners, bagging uses voting for classification. Whereas boosting refers to a family of algorithms that are able to convert weak learners to strong learners. The main principle of boosting is to fit a sequence of weak learners to weighted versions of the data. More weight is given to examples that were misclassified by earlier rounds. The predictions are then combined through a weightedmajority vote to produce the final prediction. The most widely used form of boosting algorithm is called AdaBoost, which stands for adaptive boosting. Adaboost has been used for classification purpose in this thesis.

Finally, stacking is an ensemble learning technique that combines multiple classification or regression models via a meta-classifier. The base level models are trained based on a complete training set, then the meta-model is trained on the outputs of the base level model as features.

We have assessed the performance of the aforementioned classifiers in this thesis. For all the classifiers we use MATLAB function "OptimizeHyperParameters" to look for the optimum model parameters for achieving the best classifier performance. The default mode of optimization used by the function is Bayesian Optimization (Bayesopt). The optimized parameter for each of the classifiers are: box constraint and kernel scale for SVM with a radial basis kernel, number of neighbors and distance metrics for KNN, kernel smoother type and kernel smoothing window width for NB, minimum leaf size for decision tree and number of learning cycles, learning rate and minimum leaf size for ensemble classifier with Adaboost.

### 3.14 Conclusion

In this Chapter, we developed a EMD-wavelet based EEG emotion recognition system using a statistical modeling approach. The proposed method, statistical modeling of EMD-
DWT transformed EEG signals, is presented here. The proposed system performs emotion recognition in three analysis stages: 1) Extraction of descriptive feature sets, 2) Selection of distinguishing features using ReliefF, 3) Classification across various emotion dimensions using a number of popular classification tools.
Chapter 4

SIMULATION RESULTS AND PERFORMANCE EVALUATION

4.1 Introduction

In this Chapter, the simulation result of the proposed method is evaluated in terms of recognition accuracy and F-score. Also the obtained results are compared with the state of the art emotion recognition methods. Detailed descriptions have also been carried out on the effect of ReliefF based feature selection and channel selection from neurological viewpoint.

4.2 Dataset Description

DEAP dataset [13] includes recordings from 32 participants. Each of the participants was asked to look at 40 music videos of 1 minute length and their EEG signals along with other physiological signals were recorded. The EEG signals were recorded with 32 electrodes set according to the international 10-20 system [55]. The dataset includes recordings from 12 peripheral channels, 3 unused channels and 1 status channel as well. The data from peripheral channels can also be integrated with EEG data for emotion recognition. For the purpose of this study, only EEG signals were used for classifying emotions. To label the data as corresponding to an emotion, participants used self-assessment-mannequins (SAM) at the end of each trial. For this assessment, the valence-arousal-dominance scale [56] was used. For each video, the participant rated the valence, arousal, dominance and liking on a continuous scale between 1 and 9. The original study of the DEAP database maps these scales into two level of arousal, valance, dominance and liking states. High/low states were recognized as a binary classification scheme. In this study, we follow the same scheme as the DEAP study to dichotomize the continuous
ratings into binary classes by putting a threshold in the middle of the 9 point scale.

### 4.3 Data Processing

The data set contains a preprocessed version of the original EEG signals. The original sampling rate is 512 Hz. They are downsampled to 128 Hz, EOG artifacts are removed, and a bandpass filter with cutoff frequencies of 4.0–45.0 Hz is applied. Finally, these are reorganized for classification without the hassle of processing all the data. We use this preprocessed multi-channel EEG recordings and the performed valence, arousal, dominance and liking scales in the form of high/low levels to recognize emotion state using the proposed method. The results are also compared to the method deployed in DEAP study.

### 4.4 Simulation Results

In this Section, we present the performance of the proposed method using the DEAP database. We also compare the performance of the proposed method with the results that are reported by other works in the literature using the same dataset on the same task. For classification performance measurement, accuracy is the most popular one that can identify how many samples are classified correctly. We use it here for the binary classification tasks in all emotion dimensions. Accuracy is defined as:

\[
    \text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]  

where TP, TN, FP, and FN stand for true positive, true negative, false positive and false negative events, respectively. In addition, another performance measure F1-score is provided for summarizing the results of each method in considering the balance of single number class. F1-score is measured as:

\[
    F1 - \text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

where the division of true positive outcomes by the all positive outcomes is called precision whereas recall is defined by the quotient of true positive outcomes divided by the number of positive outcomes.
The performance measurement is carried out using a Leave-one-out cross validation scheme. In each step of cross validation, one user’s feature set is used as test set while the feature set of remaining 31 users are used as training set. The obtained classification accuracies and their standard deviations have been shown in table 4.1 and 4.2. Table 4.1 contains the results obtained by using all features without any feature selection. While table 4.2 shows the effect of our feature selection scheme and the obtained improvements in classifier performance. In both the tables, accuracies achieved by utilizing both all sub bands and $\alpha, \beta$ and $\gamma$ sub bands have been reported in subsequent rows. For the sake of visualization and better understanding of the observed performance, pictorial representations have also been included in Fig. 4.1 and Fig. 4.2. Figures 4.3-4.6 show the participant wise accuracies achieved in recognition of different emotion dimensions.

Table 4.1: Classification accuracy and its standard deviation without feature selection

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Valence</th>
<th>Arousal</th>
<th>Dominance</th>
<th>Liking</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sub bands’ features+ SVM</td>
<td>72.85 ± 6.52</td>
<td>71.88 ± 6.96</td>
<td>75.59 ± 6.66</td>
<td>76.27 ± 5.77</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ SVM</td>
<td>73.53 ± 5.99</td>
<td>73.24 ± 6.14</td>
<td>74.61 ± 7.93</td>
<td>75.10 ± 6.42</td>
</tr>
<tr>
<td>All sub bands’ features+ KNN</td>
<td>72.75 ± 6.56</td>
<td>71.19 ± 7.35</td>
<td>74.80 ± 7.14</td>
<td>74.32 ± 5.93</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ KNN</td>
<td>73.05 ± 6.29</td>
<td>71.29 ± 7.16</td>
<td>74.32 ± 7.68</td>
<td>73.82 ± 6.39</td>
</tr>
<tr>
<td>All sub bands’ features+ NB</td>
<td>73.73 ± 6.73</td>
<td>72.26 ± 6.63</td>
<td>75.56 ± 7.21</td>
<td>75.78 ± 5.72</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ NB</td>
<td>74.12 ± 6.60</td>
<td>72.46 ± 6.66</td>
<td>75.39 ± 7.44</td>
<td>75.29 ± 6.61</td>
</tr>
<tr>
<td>All sub bands’ features+ Decision Tree</td>
<td>72.85 ± 6.52</td>
<td>71.88 ± 6.97</td>
<td>74.59 ± 6.66</td>
<td>74.27 ± 5.77</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ Decision Tree</td>
<td>73.53 ± 5.99</td>
<td>73.24 ± 7.14</td>
<td>74.61 ± 7.93</td>
<td>74.10 ± 6.42</td>
</tr>
<tr>
<td>All sub bands’ features+ Ensemble</td>
<td>72.27 ± 6.99</td>
<td>71.68 ± 6.83</td>
<td>74.12 ± 7.54</td>
<td>74.22 ± 7.23</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ Ensemble</td>
<td>73.05 ± 5.82</td>
<td>70.89 ± 7.99</td>
<td>74.12 ± 8.81</td>
<td>74.71 ± 6.56</td>
</tr>
</tbody>
</table>
Table 4.2: Classification accuracy and its standard deviation after feature selection With ReliefF

<table>
<thead>
<tr>
<th>Method</th>
<th>Valence</th>
<th>Arousal</th>
<th>Dominance</th>
<th>Liking</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sub bands’ features+ SVM</td>
<td>74.11 ± 6.15</td>
<td>74.02 ± 6.80</td>
<td>75.59 ± 6.77</td>
<td>75.78 ± 5.83</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ SVM</td>
<td>74.41 ± 6.71</td>
<td>74.36 ± 7.19</td>
<td>75.59 ± 6.96</td>
<td>75.10 ± 6.62</td>
</tr>
<tr>
<td>All sub bands’ features+ KNN</td>
<td>72.94 ± 6.53</td>
<td>71.77 ± 6.85</td>
<td>74.51 ± 7.62</td>
<td>74.32 ± 5.99</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ KNN</td>
<td>73.21 ± 6.56</td>
<td>71.97 ± 6.57</td>
<td>74.51 ± 7.58</td>
<td>73.93 ± 6.38</td>
</tr>
<tr>
<td>All sub bands’ features+ NB</td>
<td>75.29 ± 5.96</td>
<td>74.51 ± 6.05</td>
<td>76.86 ± 7.05</td>
<td>76.37 ± 6.30</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ NB</td>
<td><strong>75.59 ± 6.22</strong></td>
<td><strong>75.26 ± 6.69</strong></td>
<td>76.27 ± 6.63</td>
<td>75.88 ± 5.80</td>
</tr>
<tr>
<td>All sub bands’ features+ Decision Tree</td>
<td>73.53 ± 6.54</td>
<td>73.63 ± 6.49</td>
<td>74.90 ± 7.25</td>
<td>75.00 ± 6.78</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ Decision Tree</td>
<td>73.82 ± 6.72</td>
<td>73.46 ± 6.89</td>
<td>75.48 ± 7.98</td>
<td>74.90 ± 6.80</td>
</tr>
<tr>
<td>All sub bands’ features+ Ensemble</td>
<td>72.65 ± 6.50</td>
<td>72.16 ± 6.19</td>
<td>74.32 ± 7.56</td>
<td>73.82 ± 5.98</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ sub bands’ features+ Ensemble</td>
<td>73.05 ± 8.47</td>
<td>71.68 ± 6.64</td>
<td>74.02 ± 7.87</td>
<td>73.82 ± 6.49</td>
</tr>
</tbody>
</table>

Fig. 4.1: Classification accuracy and its standard deviation without feature selection
Fig. 4.2: Classification accuracy and its standard deviation after feature selection With ReliefF

Fig. 4.3: Participant wise classification accuracy in valence recognition with mean and standard deviation marked in red
Fig. 4.4: Participant wise classification accuracy in arousal recognition with mean and standard deviation marked in red

Fig. 4.5: Participant wise classification accuracy in dominance Recognition with mean and standard deviation marked in red
4.5 Comparison with Existing Works

In the following part of the study, we compare the outcome of our method with several other works available in the literature who used the DEAP database. Table 4.3 summaries the comparison of accuracies between the proposed method and the state of the art methods. It is evident that the proposed method provides better accuracy for all components of emotion compared to the existing methods listed in the Table 4.3.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Valence</th>
<th>Arousal</th>
<th>Dominance</th>
<th>Liking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Koelstra et al. [13]</td>
<td>57.6</td>
<td>62.0</td>
<td>-</td>
<td>55.5</td>
</tr>
<tr>
<td>Naser and Saha [33]</td>
<td>64.3</td>
<td>66.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shahnaz C. et al. [35]</td>
<td>64.71</td>
<td>66.51</td>
<td>66.88</td>
<td>70.52</td>
</tr>
<tr>
<td>Zhuang X. et al. [34]</td>
<td>70.9 ± 11.4</td>
<td>67.1 ± 14.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zhuang N. et al. [37]</td>
<td>69.10 ± 7.0</td>
<td>71.99 ± 7.8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mert and Akan [38]</td>
<td>72.87 ± 4.68</td>
<td>75.00 ± 7.48</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dai Y. et al. [39]</td>
<td>73.0</td>
<td>74.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>75.59 ± 6.22</td>
<td>75.26 ± 6.69</td>
<td>76.86 ± 7.05</td>
<td>76.37 ± 6.30</td>
</tr>
</tbody>
</table>

Fig. 4.6: Participant wise classification accuracy in liking recognition with mean and standard deviation marked in red
Table 4.4 summarizes the F1-scores obtained in this study and compares the results with the outcome of state of the art methods. As F1-score the geometric mean of precision and recall, the high value of F1-score strongly suggests that the proposed classification scheme both precisely and robustly in predicting high/low levels of different emotion components. additionally, from the table it is clear that in terms of F1-score, the proposed method performs way better than any other state of the art methods.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Valence</th>
<th>Arousal</th>
<th>Dominance</th>
<th>Liking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Koelstra et al. [13]</td>
<td>56.3</td>
<td>58.3</td>
<td>-</td>
<td>50.2</td>
</tr>
<tr>
<td>Shahnaz C. et al. [35]</td>
<td>74.94</td>
<td>76.68</td>
<td>76.67</td>
<td>81.94</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>81.73</td>
<td>81.14</td>
<td>83.33</td>
<td>83.43</td>
</tr>
</tbody>
</table>

**4.6 Discussion**

Some key observations can be made from the tabular and pictorial representations. Firstly, it is evident that in most of the emotion dimensions there is significant improvement in the classification performance in terms of accuracy after applying feature selection algorithm formulated using ReliefF. This shows the success of the feature selection scheme in removing the redundant features from the feature vector. Also this justifies the feasibility of using ReliefF based feature selection schemes in the application of implicit emotional tagging of multimedia contents and affective computing. Secondly, in the tables and figures, there is an interesting pattern highlighting the performance of different frequency bands in distinguishing different classes of emotion dimension. In classifying high/low valence and arousal states, the $\alpha$, $\beta$ and $\gamma$ sub bands show consistently better performance than using all sub bands. However, there is some inconsistencies in this pattern in case of dominance and liking states. This is in agreement with the fact that $\alpha$, $\beta$ and $\gamma$ sub bands of brain signals are more informative about the two major emotional dimensions, namely arousal and valence. However, there is no established ground truth about what frequency sub bands can be looked into in order to distinguish between high and low dominance or liking states. As $\alpha$, $\beta$ and $\gamma$ sub bands demonstrate a very good performance in distinguishing the most informative and descriptive emotion dimensions and also show similar or competitive distinguishing capability in dominance and liking dimensions, it can be concluded that these sub bands can act as sufficient and efficient tools to classify different
emotional states. Using only these three sub bands not only decreases the computational cost of the system to a great extent but also results in good classification performance proving its potential scope of application in studying EEG-based emotion recognition. To add to that, the classification performances also prove that the proposed channel selection method described in this thesis based on neurophysiological ground truth works well to reduce complexity while preserving good performance at the same time. This efficiency and effectiveness is highly valuable in implementation of real time practical Brain-Computer Interfaces.

The highest accuracy achieved in valence recognition is for participant 4 which is 91.25% and lowest accuracy is achieved for participant 19 which is 59.375%. For arousal recognition the highest achieved accuracy is 84.375% and it is achieved for 18 and 19. While the lowest accuracy is achieved for 6,16 and 31 which is 62.5%. In case of Dominance recognition, the highest accuracy of 87.5% is achieved for participant 17 and lowest accuracy of 56.25% is achieved for participant 28. And lastly, liking recognition for participant 2 and 22 yielded highest accuracy of 87.5% and for participant 6 and 10 the lowest accuracy of 62.5% was achieved.

The classifier which showed the highest accuracy in all the emotional dimensions is Naive-Bayes classifier. It achieved accuracies of 75.59%, 75.26%, 76.86% and 76.37% in classifying valence, arousal, dominance and liking respectively. Comparison of performances with all the other classifiers used in this study shows another remarkable outcome of the method. The proposed method is able to achieve higher accuracies by using a simpler and computationally low-cost classifier like Naive-Bayes to outperform complex classifiers like SVM, KNN etc. The performances achieved by other classifiers also show good results highlighting the robustness of the proposed feature extraction and selection scheme.

4.7 Conclusion

α, β and γ sub bands of brain wave can act as sufficient and efficient tools to classify different emotional states. This subband selection procedure along with the channel selection scheme according to neurophysiological facts about human emotions largely decrease computational cost while preserving good classification performance. At the same time, achieving higher accuracies by using a simpler and computationally low-cost classifier
like Naive-Bayes proves the robustness of the proposed features and the effectiveness of the proposed methodology.
Chapter 5

CONCLUSION

5.1 Concluding Remarks

In this thesis, a novel emotion recognition scheme is proposed which utilizes a statistical modeling based approach in the EMD-DWT domain. Proper EEG channels are first selected from a physiological and neurological viewpoint for reduced training time and computational complexity. Next, the original data is decomposed using EMD. Due to random nature of recordings of EEG data, EMD is found very effective as it is intuitive, adaptive and data-driven. After decomposing the EEG data using EMD, the three dominant intrinsic mode functions (IMFs) are selected due to their high temporal energy and frequency content. Then to obtain further discriminatory behavior and to get a better time frequency resolution, wavelet analysis on the selected IMFs is performed. Then various statistical models are fitted to the EMD-DWT decomposed EEG signal and the statistical model that best describes the data has been selected based on Bayesian information criterion. The parameters of the selected models are used to form the feature vector. ReliefF function is then used to rank the predictors with respect to their performance in correctly predicting the assigned label to a particular music video and the top ranked features are then selected to reduce redundancy. Finally, for the purpose of classification different supervised classifiers, such as SVM, KNN, NB, Decision tree and Ensemble are used on publicly available DEAP dataset.

5.2 Contributions of the thesis

The major contributions of this thesis are:

- A set of proper EEG channels are selected for data acquisition according to es-
tablished neuro-physiological facts about the correlation between different cortical regions and the elicited human emotions.

- Top three dominant IMFs are selected for signal processing due to high energy content and the correlation of human emotions and brain signals of higher frequency sub-bands.

- Different informative frequency sub-bands are investigated by choosing proper type and level of DWT coefficients.

- A feature set based on the parameters of the two best fitted statistical models is constructed for both all bands and $\alpha$, $\beta$ and $\gamma$ sub-bands' coefficient sets.

- A feature selection scheme is formulated using the ReliefF algorithm and thus removing redundant features.

- Detailed simulations have been carried out in order to investigate the performance of the proposed feature sets for the authentication and identification using EEG signals available from the DEAP database.

- The performance of our proposed method is compared with state-of-the-art methods using the same database.

- In this study significant evidences have been presented to state that the proposed method is capable of outperforming the state of the art methods in terms performance parameters, such as accuracy and F1-score.

5.3 Scopes for Future Work

However, there are still some scopes for future research, as mentioned below:

- Need for more advanced and improved algorithms remains to further reduce computational time in order to implement real-time emotion recognition systems to be applied in real-life conditions.

- Available databases other than DEAP database can be utilized for validating efficacy of our proposed method.
- Theories of deep learning and neural network algorithms can also be implemented to explore the effects in performance.
Bibliography


