EMOTION RECOGNITION USING SPATIOTEMPORAL ANALYSIS
OF ELECTROENCEPHALOGRAPHIC SIGNALS

Thesis
Submitted to
The School of Engineering of the
UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for
The Degree of
Master of Science in Electrical Engineering

By
Theus H. Aspiras
Dayton, OH
August, 2012
EMOTION RECOGNITION USING SPATIOTEMPORAL ANALYSIS
OF ELECTROENCEPHALOGRAPHIC SIGNALS

Name: Aspiras, Theus

APPROVED BY:

Tarek Taha, Ph.D.
Advisory Committee
Associate Professor, Electrical and
Computer Engineering

Eric Balster, Ph.D.
Advisory Committee
Assistant Professor, Electrical and
Computer Engineering

Vijayan K. Asari, Ph.D.
Advisory Committee Chairman
Professor, Electrical and Computer
Engineering

John G. Weber
Associate Dean
School of Engineering

Tony E. Saliba, Ph.D.
Dean, School of Engineering
& Wilke Distinguished Professor
Emotion recognition using electroencephalographic (EEG) recordings is a new area of research which focuses on recognition of emotional states of mind rather than impulsive responses. EEG recordings are found useful for the detection of emotions through monitoring the emotion characteristics of spatiotemporal variations of activations inside the brain. To distinguish between different emotions using EEG data, we need to provide specific spectral descriptors as features to quantify these spatiotemporal variations. We propose several new features, namely Normalized Root Mean Square (NRMS), Absolute Logarithm Normalized Root Mean Square (ALRMS), Logarithmic Power (LP), Normalized Logarithmic Power (NLP), and Absolute Logarithm Normalized Logarithmic Power (ALNLP) for the classification of emotions. A protocol has been established to elicit five distinct emotions: joy, sadness, disgust, fear, surprise, and neutral. EEG signals are collected using a 256-channel system, preprocessed using band-pass filters and a Laplacian Montage, and decomposed into five frequency bands using...
Discrete Wavelet Transform. The decomposed signals are transformed into different spectral descriptors and are classified using a two-layer Multilayer Perceptron (MLP) neural network. The Logarithmic Power descriptor produces the highest recognition rates, 91.82% and 94.27% recognition for two different experiments, which is more than 2% higher than when using other features.
This thesis is dedicated to:

My Lord and Savior Jesus Christ, who has brought me into the light of His salvation and has changed me through His sacrificial mercy and love.

My wife, Daisy Aspiras, who has supported me throughout our journey together and has been loving and encouraging to me ever since.

My family, who has blessed my life so abundantly with their love and sacrifice and has given me everything they could to ensure my success in life.

Christ’s love has moved me to such extremes.

His love has the first and last word in everything we do.

- 2 Corinthians 5:14 (MSG)
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>CHAPTER I: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER II: BACKGROUND AND RELATED WORK</td>
<td>4</td>
</tr>
<tr>
<td>II.1 Biological Aspects of Emotion</td>
<td>4</td>
</tr>
<tr>
<td>II.2 Psychological Aspects of Emotion</td>
<td>8</td>
</tr>
<tr>
<td>II.3 Electroencephalography</td>
<td>13</td>
</tr>
<tr>
<td>II.4 Noise Removal Preprocessing</td>
<td>14</td>
</tr>
<tr>
<td>II.5 Frequency Spectral Analysis</td>
<td>19</td>
</tr>
<tr>
<td>II.6 Current Spectral Methods in Emotion Recognition</td>
<td>22</td>
</tr>
<tr>
<td>II.7 Classification Algorithms</td>
<td>24</td>
</tr>
<tr>
<td>CHAPTER III: PROPOSED METHODOLOGY</td>
<td>28</td>
</tr>
<tr>
<td>III.1 Data Acquisition</td>
<td>29</td>
</tr>
<tr>
<td>III.2 Data Preprocessing</td>
<td>32</td>
</tr>
<tr>
<td>III.3 Feature Extraction</td>
<td>33</td>
</tr>
<tr>
<td>III.4 Classification</td>
<td>35</td>
</tr>
<tr>
<td>CHAPTER IV: RESULTS AND DISCUSSIONS</td>
<td>38</td>
</tr>
</tbody>
</table>
CHAPTER IV: EXPERIMENTAL RESULTS

IV.1 Experiment 1 ................................................................. 39
IV.2 Experiment 2 ................................................................. 40
IV.3 Histogram Analysis ....................................................... 42
IV.4 Coefficient of Variation ................................................... 44
IV.5 Nonlinear Range Compression Effectiveness ....................... 45

CHAPTER V: CONCLUSION AND FUTURE WORK ....................... 46
BIBLIOGRAPHY ...................................................................... 49
LIST OF FIGURES

II.1 A neuron in the brain made up of a cell body (soma), an axon, dendrites and terminal buttons ................................................................. 5

II.2 A neuron firing from the connection of the synapse. The terminal button contains neurotransmitters, which are released into the receptor neuron, thus creating an electrical impulse .............................................. 6

II.3 A diagram of electrical potentials when a brain neuron is firing. Once the neuron gets to the threshold potential, the neuron will shoot up in potential, and then fall quickly back to resting potential ........................................... 7

II.4 (a) James-Lange Theory (top) which says emotion comes from physiological changes, (b) Cannon-Bard Theory (middle) which says that it happens simultaneously, and (c) Lazarus Theory (bottom) which says that a cognitive assessment is needed to determine both emotion and physiological changes .......... 10

II.5 Diagrams of Plutchik's three dimensional cone and condensed wheel of emotion which show the separation of emotions for classification. .................... 12

II.6 An example of eye blinks corrupting an EEG signal through high amplitude spikes .................................................................................. 15

II.7 Frequency spectrum for: (a) 60Hz Notch Filter (top) and (b) 0.1Hz-100Hz Bandpass Filter ................................................................................ 17

II.8 A generalization of passband, rolloff, and stopband parameters for filter creation. ...... 18
II.9 Diagram of DWT decomposition of a signal into approximation and detail coefficients. ................................................................. 21

II.10 Multi-level DWT decomposition into several different coefficients. Subsampling continues throughout the process, resulting in smaller arrays of coefficients. ............. 21

II.11 Two-layer perceptron network with two sets of weights, w and v. ......................... 25

III.1 A flow chart of the Brain-Computer Interface used for our experiments. After acquiring the EEG data, the data is processed through three other stages (Data Preprocessing, Feature Extraction, and Classification) to provide our recognition results. ................................................................. 29

III.2 EGI's Geodesic EEG 300 System with the 256-channel electrode sensor net applied to a subject. ................................................................. 30

III.3 Electrode layout of the 256-Channel EEG System from EGI, Inc. ......................... 31

III.4 The sample visual stimuli used from the International Affective Picture System. The left picture shows fear while the right picture shows sadness. ....................... 32

III.5 A nonlinear boundary created to separate a set of data. When a linear boundary is formed, some data may be classified incorrectly. ................................. 37

IV.1 Histogram distributions of RMS (left column), ALRMS (middle column), and LP (right column) for different frequency bands. ................................. 43

V.1 EEG signal of an eye electrode using eye artifact removal techniques. (a) Original EEG Signal, (b) Corrected EEG Signal ................................. 47
LIST OF TABLES

II.1 Decomposed signal and their respective frequency bands by DWT, and the standard frequency bands ................................................................. 22

IV.1 Recognition rates for classical, normalized, and logarithmic power feature metrics for Experiment 1. Logarithmic power gave the highest recognition of 91.82% across all emotions................................................................. 40

IV.2 Normalized and logarithmic power feature metric recognition rates for Experiment 2. The logarithmic power feature metric provided the highest average recognition rate of 94.27% across all emotions................................. 41

IV.3 The coefficient of variation for the different feature metrics across different frequency bands. The normalized and compression methods (REE, NRMS, ALREE, ALRMS, and LP) gave the lowest coefficients of variation, which correlates to the higher recognition rates found by using these methods. .............. 44
CHAPTER I
INTRODUCTION

Emotion Recognition is a new area of research that seeks to find metrics to define specific emotional states. This research has several motivations that would prove useful in real-world applications. In the medical field, a brain computer interface (BCI) can be created to have communication between patients and doctors, which can be an invaluable technology specifically for disabled patients and patients incapable of communicating through speech and traditional methods. This could even give the ability to diagnose different diseases through the deviation of electroencephalography (EEG) activations in the brain. With the area of communication, a BCI can be developed to control specific activities with brain activity. Again, patients with disabilities can benefit from this technology by being able to control functions, like moving a cursor on a screen or navigating with an automated wheelchair, which would otherwise be impossible with physical movement. For vigilance tasks, detection of negative intentions and motives can be done during an interrogation of a subject. Another area of study used for emotion recognition would be the quantification of brain efficiency during tasks. This would be useful for businesses wanting specific stimuli which would increase brain efficiency during work tasks. This can also be used for spatial disorientation/stress detection for fighter pilots, which would improve their capabilities during intense combat.

There are several challenges that are found in developing a BCI for emotion recognition. For our experiments, we needed to find a method that would provide the best elicitation of
different emotions. This elicitation should contain the optimal stimuli for a specific emotion, which defines a specific strength/range of elicitation. Also when the elicitation has happened, we need to validate that they went through that specific emotion. This leads to uniqueness in responses for various people. Each person can be given the same stimuli, but can lead to a different reaction and thought process, which would be hard to validate for different people. Therefore, if we use a well-known stimuli set for elicitation of emotions, we can then generalize most reactions to the stimuli as the same response. Another set of challenges is the noise found in EEG. EEG noise can come from a variety of sources, like electrical noise, eye blinks/movement, and bad EEG channels. We must develop methods and techniques to remove these noises for a clean EEG that can be used for finding features. Another challenge is developing methods that can be used in real-time. Due to the high volume of data that is collected through the EEG system, we must find algorithms that would decrease processing time and improve efficiency of the system.

Therefore, we have three different research objectives for our emotion recognition research. We first want to develop methods to incite emotional responses in detecting brainwave features through electroencephalography recordings. We then want to optimize various noise-reduction, feature extraction, and classification algorithms. Finally, we want to create a real-time BCI, compiling all methods and algorithms for emotion detection and recognition.

This paper defines several algorithms and techniques to create a BCI for emotion recognition. We propose an array of algorithms to use for all aspects of the BCI: preprocessing, feature extraction, and classification. We also want to improve on current technology, specifically in feature extraction, to allow better results in classification. For preprocessing, we propose using two different types of frequency filters, a 60Hz line filter and a 0.1Hz-100Hz Bandpass filter, and Surface Laplacian montage. For feature extraction, we use the discrete
wavelet transform (DWT) to do spectral analysis on the data and derive a new feature called Logarithmic Power which represents the data in the best form possible for use in the classifier. For classification, we use the multilayer perceptron network (MLP).

To illustrate the performance of the system, we collect data from a 256-Channel high density sensor net, which provides unparalleled resolution of EEG scalp activations, and apply this data into our brain computer interface. We do two different tests based on the training type of our classifier to test the performance of our proposed feature extraction technique against other known feature extraction techniques.

The rest of the thesis is structured as follows: After introducing background knowledge on the biological and psychological aspects of emotion, historical EEG research, and current methods on EEG noise reduction, feature extraction, and classification in Chapter II, we present our improved brain computer interface with an optimized feature metric in Chapter III. In Chapter IV, we show the testing performance of the system and compare the performance of our new feature metric, Logarithmic Power, against conventional feature metrics and state-of-the-art feature metrics. We then conclude and display our future work in Chapter V.
CHAPTER II
BACKGROUND AND RELATED WORK

In this chapter, we will explore the psychological aspects of emotion, go over a brief history of EEG research, define currently state of the art algorithms found in EEG research, and then look at specific algorithms pertaining to EEG Emotion Recognition.

II.1 BIOLOGICAL ASPECTS OF EMOTION

The human brain is the center of the nervous system, which controls and monitors actions and reactions within our body. These include muscular activity, organ function, memory retention, chemical changes, emotions, awareness, etc. The brain contains 50-100 billion neurons that work together inside the brain to perform these functions. Each of the neurons inside your brain can have up to 10,000 synaptic connections, which are junctions that connect neurons to other neurons. The interweaving of neurons is what promotes the building of memory and learning capabilities.

Each neuron, as shown in Figure II.1 is comprised of three parts: the cell body, dendrites, and axons. The cell body, also called the soma, is the center of the neuron which contains the nucleus of the cell and controls protein synthesis. The dendrites are the cellular extension from other neurons which provide the input from a variety of sources, including signals from other neurons, muscles, sensory organs like the eyes and ears, etc. The axons are the output of the
neuron. Each neuron would contain only one axon, but can branch out through terminal buttons, which are extensions from the neuron to create synaptic junctions with other neurons.

![Diagram of neuron components](image)

**Figure II.1:** A neuron in the brain made up of a cell body (soma), an axon, dendrites and terminal buttons.

The neurons have ion channels which aid in the propagation of electrical impulses, also called action potentials. These ion channels would be found in the membrane of the soma and the axon. The ion types used in the brain are Sodium, Potassium, Chloride, and Calcium, which are necessary for the firing of the neurons. Also necessary for the firing of the neurons is neuron stimulation. This can come from many different sources like pressure, chemical transmitters, and changes of electrical potential. These stimuli would cause an ion channel to open, thus allowing the flow of ions in the neuron and changing the potential of the neuron. This change in potential is also called neuron firing.

Neuron firing is the method by which neurons communicate. Each of the synapses in the brain is created by the connections of terminal branches from the output (presynaptic) neuron to dendrites in the input (postsynaptic) neuron. The connections for the synapses can be either
chemical or electrical. The action potential of the presynaptic neuron travels to the synapses, depolarizing the neuron and opening an ion channel. This would allow ions to flow to the terminal branch ends, releasing a neurotransmitter which binds to the postsynaptic neuron. The reception of these neurotransmitters would open up the ion channels in the postsynaptic neuron creating two types of activations: depolarization, which increase the action potential, and hyperpolarization, which decreases the action potential. Figure II.2 shows a neuron, labeled with its different parts, firing using the synapses and neurotransmitters.

Figure II.2: A neuron firing from the connection of the synapse. The terminal button contains neurotransmitters, which are released into the receptor neuron, thus creating an electrical impulse.
Action potentials for a neuron are the electrical impulses that flow through from neuron to neuron. The resting potential for a neuron is -70 mV and the threshold potential, meaning the potential for the neuron to fire, is -55 mV. When the neuron depolarizes, it brings the neuron closer to the threshold potential. When the neuron hyperpolarizes, it brings the neuron down to the resting potential. Once the neuron goes above the threshold potential, the neuron activates, thus thrusting the activation potential upwards, reaching as high as +100 mV, and then shoots back to resting potential. The progression of the action potential is demonstrated by Figure II.3.

![Figure II.3: A diagram of electrical potentials when a brain neuron is firing. Once the neuron gets to the threshold potential, the neuron will shoot up in potential, and then fall quickly back to resting potential.](image)

Neurons that are interconnected communicate to form different functions and tasks. Thus, the brain can be grouped into different structures called lobes. The brain contains 4 types of lobes: the frontal lobe, parietal lobe, occipital lobe, and the temporal lobe. The frontal lobe does decision-making and complex mental functions, which is also tied to emotional states. The parietal lobe performs spatial awareness tasks and processes sensory communication from the
body. The occipital lobe is highly connected to visual tasks and the temporal lobe develops auditory perception and also remembers faces/scenes. Even deeper inside the brain is the Limbic system, which controls different autonomous functions in the body, like heart-rate and blood-pressure. One specific portion of the limbic system that is of interest is the amygdala, which helps to process and remember emotional reaction and is highly connected to the pre-frontal cortex.

II.2 PSYCHOLOGICAL ASPECTS OF EMOTION

As different portions of our brain work together to perform different functions and memories, these thought processes become cognition. Cognition is a collection of thought processes that desire acquisition and application of knowledge, like awareness, perception, intuition, and memory. Traditionally, emotion was not classified as a cognitive process, due to the definition. Thomas Aquinas separated psychology to how we know and how our knowledge affects us. In most experiments on cognition, the influence of emotion on cognitive states was ignored, therefore assuming a neutral emotional state. For example, how might a person in a sad mood have their cognitive processes affected? The focus of attention may not be the task at hand, but at the memories of what influenced the sad emotion. We must then find the connection and interaction between cognition and emotion.

Emotions have a heavy influence on our thought and cognitive processes, so to analyze different emotions, we must look at the relationship between emotion and cognition and how our brain regulates emotions. Clore and Palmer [1] show the relationship between emotion and judgments and how values placed on certain judgments are influenced by these emotions. The development of emotions in our brain is another important topic to discuss. James-Lange's theory, depicted by Figure II.4.a, suggests that emotions come from physiological changes [2],
which are also shown by Naqvi et al. [3] through the somatic-maker hypothesis. A good example of this is when people force a smile upon their faces when they are sad, they start to become emotionally happy from the physiological changes. Cannon-Bard's theory, depicted by Figure II.4.b, shows that emotions occur together with physiological changes [4]. An example of this is a person shaking and having fear simultaneously due to a frightening event. In Folkman and Lazarus [5] the cognitive assessment of a certain thought would induce physiological changes and emotion simultaneously, which is shown by Figure II.4.c.
When classifying different emotions, we must note the duration of some emotions. Some emotions last for only a few moments, like surprise. Other emotions can last a lifetime, like love. Depending on the situation, emotions can have a varying duration, with different strengths of emotion. To define the separability and strengths of different emotions, Plutchik [6] defined 8 primary emotions with different intensities and mixtures, which forms new types of emotion. To
create a visual emotional space, Plutchik devised a three-dimensional cone and a wheel which describes the relationship of various emotions, as shown in Figure II.5. Another psychological process that should be looked at is emotion regulation, which is the process of inhibiting emotions. This could be the suppressing of physiological signals, thoughts, cognitive states, and even behaviors. One example of this is holding back tears in the death of a loved one. This process is dependent on the person and may vary with different types of situations and stimuli.
Figure II.5: Diagrams of Plutchik's three dimensional cone and condensed wheel of emotion which show the separation of emotions for classification.
II.3 ELECTROENCEPHALOGRAPHY

Electroencephalography (EEG) is the time-domain recordings of electrical potentials across the scalp. As the brain processes different thoughts and emotions, the neurons within the brain fire thus creating electrical impulses within the brain. These electrical impulses then radiate towards the scalp of the brain for electrode sensors to pick up. Several of these electrodes can be placed across the surface of the scalp, up to 256 electrode channels, to obtain a high resolution recording. These electrical impulses are detected as small voltage differences on the scalp, ranging from 10μV to 100μV, and must be amplified for analysis. Sensors across the scalp can have high sampling rates, ranging from 256-512Hz for clinical studies and up to 20kHz for different studies.

Through the development of another technology called functional Magnetic Resonance Imaging (fMRI), there have been many debates as to which is the best mode for brainwave analysis. The fMRI has a very high resolution, spatial three-dimensional representation of brain activations (up to 0.1mm) but has poor temporal resolution. The EEG has fairly poor spatial resolution and is a scalp surface representation, but has very high temporal resolution (up to 20kHz). The EEG is also a direct measurement of neuronal brain activations radiating to the scalp, while the fMRI measures hemodynamic response (blood flow), which occurs roughly 5 seconds after brain activation. Therefore, the difference between the modes is localization vs. interaction. Localization finds a specific region of the brain that activates during a specific set of stimuli while interaction describes activation dynamics in the brain, showing activation changes overtime. EEG can offer low-resolution source imaging while exploiting these activation changes.

Therefore, EEGs are used for spatiotemporal analysis of brainwave signals. We can use the EEG to find activations from different parts of the brain, like the frontal lobe and the limbic
system, and find combinations of these activations for a specific response. We can also map brain signal paths due to the high temporal resolution of the EEG and see the instantaneous activation results from affective stimuli. Through the affective stimuli, we can define several emotional brain states, which include emotion duration, emotional arousal strength, and even a combination of emotions.

Electroencephalographic recordings have been used since 1924 to classify several types of psychological phenomena and diagnose different types of medical disorders. The P300 response, found by Chapman and Bragdon [7], is one of the most heavily researched areas that show the relationship between the event-related potential inside the brain to a visual stimulus. This particular relationship sparked several other research topics, like Farwell [8] using the P300 response with the Guilt Knowledge Test for Interrogative Polygraphs, and Wall and Ehlers [9] determining the effects of alcohol on the P300 response of Asians. Since EEG recordings are quantified in several voltages detected on the scalp of the brain, several efforts have been made to find suitable transformations for the analysis of the data.

II.4 NOISE REMOVAL PREPROCESSING

With any electroencephalographic recording, there will be high amounts of noise present in several areas. One type of noise that is present in EEG is 60Hz line noise, which is created by external noise from electrical equipment near the testing sight. This noise is usually in a narrow frequency band and may affect some or all electrodes. Another type is eye blinks, which are very common in EEG and are characterized by high amplitude events in the electrodes near the eyes. These particularly disrupt the display of frontal activations due to its high amplitude. Figure II.6 shows an example of eye blinks corrupting EEG data. Another type is eye movements, which are
due to the reorientation of the eye. These spread more readily in the EEG and happen in small intervals.

![Figure II.6: An example of eye blinks corrupting an EEG signal through high amplitude spikes.](image)

Noise/artifact removal is an important area of study for Electroencephalography (EEG). While EEG electrodes detect diminutive activations in the brain, they are susceptible to a wide array of noises. A comparison between the recorded EEG and the electrocardiogram (ECG) by Dirlich et al. [10], who show that cardiac field artifacts are high amplitude potentials which affect EEG performance. Dewan et al. [11] are able to remove these ECG-type artifacts by developing a noise model based on energy functions to subtract the noise from the recorded EEG. Muscle activations, such as jaw clenching and facial movements, are also potential sources of artifacts in the EEG. Narasimhan and Dutt [12] finds that muscle artifacts hidden in EEG potentials can be removed by least mean squared adaptive predictive filtering. De Clercq et al. [13] finds that using a blind source separation technique called Canonical Correlation Analysis proved to be better for muscular artifact removal than low pass filters and Independent Component Analysis.
Ferdjallah and Barr [14] develop different types of adaptive FIR and IIR notch filters to remove power line noise in EEG signals.

The noise/artifact removal research is most prevalent in the removal of eye blink/movement artifacts. One method by Jervis et al. [15] finds the cross-correlation between the Electroculography (EOG) recordings and the EEG and subtracts a fraction of the EOG signals from the recorded EEG to yield a cleaned EEG. Ramanan et al. [16] uses Haar wavelets to distinguish between eye artifacts and clean EEG, offering an algorithm that can detect and remove artifacts in epileptic EEG. Neural networks have been used by Erfanian and Mahmoudi [17] to suppress eye artifacts in the EEG as with many others [18-19]. The techniques that are most widely used are the Blind Source Separation (BSS) techniques. Gomez-Herrero et al. [20] uses BSS to develop an automated artifact removal system without the use of an EOG signal. Liu and Yao [21] used Principal Component Analysis to correctly identify EOG sources and correct the EEG from them. From the different types of blind source separation techniques used, Independent Component Analysis has been used extensively to remove ocular artifacts in EEG. Zhou et al. [22] uses ICA to remove both eye artifacts and power line noise. Several others [23-26] also use Independent Component Analysis to isolate eye artifacts in the EEG and remove them to create a clean EEG.

For usage in our experiments, the most prominent needs for the system would be a line-noise removal algorithm, a bias removal algorithm, and an eye-blink removal algorithm. For the development of the line noise removal algorithm, we implement a notch filter. The filter passes all frequencies except frequencies located near the stopband, which is centered on a specific frequency. For our experiments, we design a notch filter at 60Hz to remove noise in that selected frequency, which is shown by Figure II.7.a. To remove the biases in the system and obtain a frequency range that is needed for feature extraction, a 0.1Hz to 100Hz Bandpass filter is created,
which is a filter that passes all frequencies located within a certain frequency band. When defining these filters, we provide a passband gain, which controls the amount of the signal that remains in the passband, stopband gain, which controls the amount of the signal that the filter attenuates in the stopband, and a rolloff frequency range, which controls the transition frequency band from passband to stopband and vice versa. Figure II.8 shows a diagram of how the frequency filters are designed.

Figure II.7: Frequency spectrum for: (a) 60Hz Notch Filter (top) and (b) 0.1Hz-100Hz Bandpass Filter
Surface Laplacian Montage

Different montages can be applied to a set of EEG signals to allow normalization of data and to mitigate artifacts that arise often in EEG. One of the most popular montages named Average Reference allows the data to have an extra channel, which is the average of all the channels. Instead of having an extra channel, Surface Laplacian montage subtracts the weighted average of surrounding electrodes from each electrode. Law et al. [27] finds that Surface Laplacian Montage reduces blink and eye source contamination. Equation II.1 shows the modifications of each channel with respect to other channels.

\[
x_{i}^{\text{new}} = x_i - \frac{1}{N} \sum_{j=1}^{N} w_{ij} x_j
\]

where \( x_i \) is the signal magnitude of the \( i \)th channel where \( i = 1, 2, \ldots, N \), \( x_j \) are the signal magnitudes of surrounding channels, \( w_{ij} \) is the weight matrix for each channel corresponding to
every other channel, and N is the number of channels. This technique was used by Nunez and Westdorp [28] to prove that the surface Laplacian montage estimates cortical surface potentials for raw EEG data. By reducing the commonality between EEG electrodes and using neighboring electrodes to influence the electrode of interest, we can further reduce the influence of noise and artifacts that are present in EEG recordings. To remove the mean value of all the electrodes, we will assume that the weight matrix is filled with the value of 1.

II.5 FREQUENCY SPECTRAL ANALYSIS

Frequency analysis, one of the more popular transformations of EEG data, specifically transforms the data into 5 specific frequency bands (Delta, Theta, Alpha, Beta, and Gamma). Armitage et al. [29] and Feinberg et al. [30] look at the Delta frequencies (0 - 4Hz) in several sleep studies for depressed adolescent females and sleep-loss, respectively. Schacter [31] analyzes the impact of the Theta frequency band (4 - 8Hz) for different psychological phenomena. The Alpha frequency band (8 - 13Hz) is found useful by Itil [32] for the EEG of adult schizophrenics and by Nowak and Marczynski [33] for trait anxiety from stress. Rangaswamy et al. [34] look at the Beta frequencies (13 - 30Hz) to determine the magnitude at different spatial locations and its connection to an imbalance of the central nervous system in alcoholics. The correlation between Gamma frequency bands (30Hz - 100Hz) and visual stimuli are found by Müller [35], who determines that Gamma frequencies are induced by attention to visual stimuli and response to emotional pictures.

Since many features can be found using frequency spectral analysis, the Discrete Wavelet Transform (DWT) has been a popular technique for the decomposition of EEG data. Kumar et al. [36] uses DWT to remove ocular artifacts in the EEG without using an Electrooculography (EOG) signal. Others, like Whittaker and Siegfried [37] and Sherwood and Derakshani [38], use
the discrete wavelet transform for analysis of EEG data on visual stimuli and classification of different motor tasks and cognitive tasks, respectively.

**Discrete Wavelet Transform**

The wavelet transform is a powerful tool which transforms a signal into a series of wavelets. This decomposition is done by dilating the mother wavelet into a set of coefficients and repeating the decomposition to obtain the series of wavelets. Wavelets offer joint time-frequency analysis upon the signals through the derivation of the coefficients. For each derivation, coefficients in each wavelet represent the signal in their specific frequency band. Also, we can specify a time window in which the wavelet is calculated to produce several sets of wavelets over time. Finding what time interval and how many bands to decompose the signal depends on the types of features we are looking for. We can use the information given based on the specific frequency bands and time scales for our features.

To make calculations of wavelets faster and to utilize the discrete nature of our sampled data, we can apply the discrete wavelet transform, which is calculated by passing the signal through a set of filters. We can first apply a low-pass filter with impulse response $h[n]$ which is modeled from the lower frequency range of the signal, as shown in Equation II.2.

$$y[n] = x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]$$  \hspace{1cm} (II.2)

where $x[n]$ is the original signal, $y[n]$ is the resulting signal, and $*$ is the convolution operator.

A high-pass filter with impulse response $g[n]$ is also applied, resulting in two signals with half of the original frequency band. Because of this, according to Nyquist theorem, half of the samples are to be discarded for each of the signals, as shown in Equations II.3 and II.4 respectively.

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k]$$  \hspace{1cm} (II.3)
\[
y_{\text{high}}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n + 1 - k]
\]  

(II.4)

Those signals corresponding to the higher frequencies are detail coefficients and those corresponding to the lower frequencies are approximation coefficients as illustrated in Figure II.9.

Figure II.9: Diagram of DWT decomposition of a signal into approximation and detail coefficients.

Figure II.10: Multi-level DWT decomposition into several different coefficients. Subsampling continues throughout the process, resulting in smaller arrays of coefficients.

By applying several cascading filters to the resulting approximation coefficients, the subsampled data can be further decomposed into sets of wavelets as shown in Figure II.10. Finding the wavelet that is appropriate for extracting features is important. For fast computation, we used Daubechies wavelet of order 2. Since we can choose how many levels of wavelets to apply to our data, we chose to decompose the signal into five frequency bands often used in EEG literature which are shown in Table II.1
<table>
<thead>
<tr>
<th>Decomposed Signal</th>
<th>Frequency Bands (Hz)</th>
<th>Standard Frequency Bands (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma (D2)</td>
<td>31.25 – 62.5</td>
<td>30.0 - 100.0</td>
</tr>
<tr>
<td>Beta (D3)</td>
<td>15.625 – 31.25</td>
<td>13.0 - 30.0</td>
</tr>
<tr>
<td>Alpha (D4)</td>
<td>7.8125 – 15.625</td>
<td>8.0 - 13.0</td>
</tr>
<tr>
<td>Theta (D5)</td>
<td>3.90625 – 7.8125</td>
<td>4.0 - 8.0</td>
</tr>
<tr>
<td>Delta (A5)</td>
<td>0 – 3.90625</td>
<td>0 - 4.0</td>
</tr>
</tbody>
</table>

Table II.1: Decomposed signal and their respective frequency bands by DWT, and the standard frequency bands.

II.6 CURRENT SPECTRAL METHODS IN EMOTION RECOGNITION

The recognition of emotion using EEG data has been an ongoing research topic for several researchers. Tran and Asari [39] used Independent Component Analysis (ICA) and Multi-Layer Perceptron (MLP) networks to recognize four different mental tasks (rest, imagined movement, visual spelling, and solving a math problem). The combination of these techniques resulted in 69% to 98% recognition across all of the mental tasks. Takahashi [40] uses video as an emotional stimulus and calculated several different features, like mean, which resulted in about 41% recognition of emotion. To incorporate frequency analysis in the recognition of emotion, Kostyunina and Kulikov [41] analyzes different frequency bands of the EEG and found that the Alpha frequency band contains enough information to differentiate emotional states. Ray and Cole [42] show that instead of the Alpha frequency band containing information of emotion, it is found in the Beta frequency band while the Alpha frequency band reflects attention. Ko et al. [43] used relative power values as features and had 70% recognition for some emotions and 10% recognition for other emotions. Murugappan et al. [44] proposes Recoursing Energy Efficiency (REE) and a modified algorithm, the Absolute Logarithmic Recoursing Energy Efficiency [45],
which has classification rates of 83%, 80%, and 73% for 64, 24, and 8 channels of EEG data, respectively.

**Recoursing Energy Efficiency**

After utilizing the wavelet decomposition technique to divide the signal into several frequency bands, we calculate different feature metrics for each of the frequency bands. Some of the classical features used in the experiments are minimum, maximum, mean, standard deviation, energy, power, and root mean square (RMS). Murugappan et al. proposes using energy computed from the wavelet coefficients in the techniques Recoursing Energy Efficiency (REE) and Absolute Logarithm Recoursing Energy Efficiency (ALREE). This technique offers a normalization of the energy across each of the frequency bands. To show how Recoursing Energy Efficiency works, we first compute the energy in each frequency, as shown in Equation II.5

\[
E_f = \sum_{n=1}^{N} (x_{n_f})^2
\]  

(II.5)

We then find the total energy in each of the frequency bands, as shown in Equation II.6.

\[
E_{total} = \sum E_f
\]  

(II.6)

We then compute the Recoursing Energy Efficiency, which is the percentage of each frequency band's energy to the total energy, as shown in Equation II.7.

\[
REE_f = \frac{E_f}{E_{total}}
\]  

(II.7)

The Absolute Logarithm Recoursing Energy Efficiency (ALREE) uses the REE value, but computes the absolute logarithm of that value, which is shown in Equation II.8.

\[
ALREE_f = \left| \log \left( \frac{E_f}{E_{total}} \right) \right|
\]  

(II.8)
II.7 CLASSIFICATION ALGORITHMS

After developing a method to obtain features from the data, we need to find a suitable classifier to classify the data for our needs. Murugappan et al. used K-Nearest Neighbors (KNN) and Linear Discriminant Analysis (LDA) to classify the relevant features. The problem with K-Nearest Neighbors is that it is an unsupervised learning method, which may group a feature vector of a particular class to another class. The problem with Linear Discriminant Analysis is that it is linear, meaning that feature data that is lying on a potentially nonlinear manifold will not be properly classified. The best type of classifier that would be good for classifying emotions would be one that has nonlinear manifold learning capabilities.

Multilayer Perceptron Network

These networks are known for their ability to learn nonlinear manifolds with different sizes of training sets and generalize data with fast operation. What gives the MLP network such powerful generalization and nonlinearity is the ability to use hidden nodes. These hidden nodes allow a modification of weights that are nonlinear, due to the threshold function of each node. Several layers can be introduced to provide deeper learning of the training data.

The multilayer perceptron network has N inputs, L hidden nodes, M outputs, and is connected by a set of weights v and w as shown in Figure II.11.
The output of each node in the system is computed by first scaling the inputs by the synaptic weights and summing the scaled values and then applying a sigmoid threshold function as in Equations II.17 and II.18.

\[
Net_j = \sum_{k=1}^{L} w_{jk} u_k + w_{j0}
\]

\[
y_j = f(Net_j) = \frac{1}{1 + e^{-Net_j}}
\]

where \( w_{jk} \) are the weights between neurons, \( w_{j0} \) are the weight biases, \( Net_j \) is the sum of all the weights and inputs, \( f(Net_j) \) is the threshold function, and \( y_j \) is the output of the layer. The weight updates for all layers are given by Equation II.19

\[
w_{jk}(t+1) = w_{jk}(t) + \Delta w_{jk}(t)
\]
where $w_{jk}(t)$ are the old weights, $\Delta w_{jk}(t)$ are the changes in weight, and $w_{jk}(t+1)$ are the new weights. The change of weight in each iteration is computed by a gradient descent method, also known as the steepest descent method, which tries to find the direction that decreases the error function. The following function gives the equation for the gradient descent algorithm.

$$ x_{k+1} = x_k + \alpha_k d_k $$

where the search direction is the negative gradient $d_k = -\nabla f(x_k)$ and $\alpha_k = \eta$ which is a step size, also known as learning rate parameter. The negative gradient can also be considered as the slope of a multi-valued function. So by finding the slope that maximizes the gradient descent, we can speed up the minimization of the error function. In the case for the back-propagation neural network, we must find weights that minimize the error function. To derive the weights, we start with the mean-squared error equation between the desired output and the actual output as

$$ E = \frac{1}{M} \sum_{j=1}^{M} (d_j - y_j)^2 $$

where $d_j$ is the target output and $y_j$ is the actual output for all $M$ patterns. $\Delta w_{jk}(t)$ is proportional to the negative of the derivative of the error with respect to the weights. By taking this derivative, we can find the weights that minimize the error. By using chain rule, we can derive more functions to compute the weights. We will first compute the weights for the last layer. The derivative of the error is given by Equation II.22.

$$ \frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial \text{Net}_j} \cdot \frac{\partial \text{Net}_j}{\partial w_{jk}} $$

By continuously applying the chain rule across this equation, we can derive the weights of the hidden layer nodes and the output layer nodes as in Equations II.23 and II.24 respectively.

$$ \Delta v_{ki}(t) = \eta \cdot \sum_{j=1}^{M} (d_j - y_j) \cdot f'(\text{Net}_j) \cdot w_{jk} \cdot f''(\text{Net}_k) \cdot x_i $$
$$\Delta w_{jk}(t) = \eta \cdot (d_j - y_j) \cdot f'(Net_j) \cdot u_k$$  \hspace{1cm} (II.24)$$

where $\Delta w_{jk}(t)$ are the weights of the output layer, $x_j$ are inputs to the hidden layer, $Net_k$ is the sum from the input nodes of the hidden layer, $\Delta w_{jk}$ are the weights of the output layer, $Net_j$ is the sum from the inputs of the output layer, $u_k$ are the inputs to the output layer, and $\eta$ is the learning rate parameter such that $0 < \eta < 1$.

The use of the Recoursing Energy Efficiency algorithm for emotion recognition is verified [46]. To improve [47], we proposed the Normalized Root Mean Square (NRMS) and the Absolute Logarithmic Normalized Root Mean Square (ALRMS), which provides higher recognition than the REE and ALREE algorithms. We propose using a different technique called Logarithmic Power (LP), which compresses the range of values used for classification, and test its effectiveness against all previous algorithms.
CHAPTER III
PROPOSED METHODOLOGY

To analyze the EEG data for emotion recognition, a brain computer interface is developed to first acquire the EEG data needed from each subject. Once the data is obtained, the data is preprocessed to filter out noise and artifacts in the EEG using two frequency filters and the Laplacian Montage. In the feature extraction stage, DWT based spectral decomposition and spectral descriptors are used to find significant features that can be useful for our emotion recognition process. These features are then sent to a Multilayer Perceptron Neural Network classifier to determine the relevance of the features and the recognition of emotion through these features. Figure III.1 shows each process done in the brain-computer interface.
Figure III.1: A flow chart of the Brain-Computer Interface used for our experiments. After acquiring the EEG data, the data is processed through three other stages (Data Preprocessing, Feature Extraction, and Classification) to provide our recognition results.

### III.1 DATA ACQUISITION

The data acquisition system used is the Geodesic EEG 300 System (GES) by Electrical Geodesic, Inc., which has a 256-channel high-density HydroCel Geodesic Sensor Net (HCGSN) for high resolution EEG data. The system has several analysis tools and algorithms for use on the EEG. The E-Prime System by Psychology Software Tools, Inc. is also used to display our stimuli for our subjects. Figure III.2 shows the 256-channel high-density HCGSN applied to a subject and Figure III.3 shows the electrode layout of the system.
Figure III.2: EGI's Geodesic EEG 300 System with the 256-channel electrode sensor net applied to a subject.
To conduct our study, five subjects are taken, all university students, and presented twenty-five images, five images for each of the five different emotions (Joy, Disgust, Sadness, Fear, and Neutral). Images are taken from the International Affective Picture System [48], which is a standardized set of images to invoke emotions from human subjects. These images are displayed randomly for duration of 30 seconds for each person, with a rest period of 10 seconds in between each image. We use the 256-Channel EEG data acquisition system from EGI, Inc. to obtain our recordings and the E-Prime system from PST, Inc. to present visual stimuli, which are
shown in Figure III.4. The EEG data sampling rate used for this study is 250Hz, which provides enough samples for the analysis of our EEG data.

Figure III.4: The sample visual stimuli used from the International Affective Picture System. The left picture shows fear while the right picture shows sadness.

### III.2 DATA PREPROCESSING

Several preprocessing algorithms have been applied to the raw EEG data to eliminate noise/artifacts in the signals. Different frequency filters are applied to remove the unwanted noise from the signals and to allow a specific frequency band of interest. A band-pass filter from 0.1Hz to about 100Hz is applied to get a range of frequencies and to remove biases that may be present in the signal. A 60Hz notch filter is also applied to remove 60Hz line noise that was
present in the signal. A Surface Laplacian Montage is then implemented to mitigate most eye blink/movement noise and artifacts. By defining a specific range of frequencies and applying a montage to reduce the amount of noise that is present in the EEG, the data is optimized for better extraction of necessary features in the data. The passband gain is chosen to be 99% of the original, the stopband gain to be 1% of the original, and the rolloff frequency range to be 2Hz.

III.3 FEATURE EXTRACTION

There are 2 steps used for our feature extraction. We first compute the Discrete Wavelet Transform (DWT) and then compute the proper feature metrics.

Proposed Feature Metrics

The proposed techniques are modeled by the Recoursing Energy Efficiency (REE) feature extraction method, since it obtained fairly high recognition rates. In the REE algorithm, energy is used to calculate the normalization of the features across the different frequency bands. It is found that by first using the wavelet transform to transform the data into the frequency bands, subsampling occurs, which skews the energy calculation. Since there is subsampling in the DWT algorithm, the calculation of energy does not account for the number of samples in each frequency band, thus compression occurs in the signal, especially in lower frequency bands. This means that since the energy for the lower frequency band is calculated, there is smaller number of samples than if we had a higher frequency band. This also means that when calculating energy, energy for higher frequencies will be higher because of the lack of normalization of the energy. Therefore, it is decided to use a measure that normalized the feature based on the number of samples, like mean or power.

Even though these conventional features normalize the data, it is found that it did not work well because the subsampling of the wavelet function in conjunction with the calculation of
energy allows a sort of compression of the data. The fewer samples you have, the more you compress the data so that you will obtain a smaller value. In the proposed method, the root mean square value is used rather than the energy values because of the compression ability of the value. Since compression of the energy inside the normalization technique provides better recognition, RMS would provide even more compression. Equation III.9 shows the RMS value of each frequency band, Equation III.10 shows RMS across all frequency bands, and Equation III.11 shows the normalized root mean squared value (NRMS)

\[
RMS_f = \sqrt{\frac{\sum_{n=1}^{N} (x_{nf})^2}{N}} \tag{III.9}
\]

\[
RMS_{total} = \sum RMS_f \tag{III.10}
\]

\[
NRMS_f = \frac{RMS_f}{RMS_{total}} \tag{III.11}
\]

Since the Recoursing Energy Efficiency algorithm is being modeled, the root mean square value can be applied to the modified version, Absolute Logarithm Recoursing Energy Efficiency. Equation III.12 shows the absolute logarithm normalized root mean squared value (ALRMS).

\[
ALRMS_f = \left| \log \left( \frac{RMS_f}{RMS_{total}} \right) \right| \tag{III.12}
\]

**Logarithmic Power Representation**

Since compression is utilized from within the normalization technique of Recoursing Energy Efficiency, boosting the compression from within is proposed, which would provide better results, even without the usage of the normalization. This is due to the high power differences with eye blinks and certain activations in the brain. By compressing these values,
greater separation of the features happens without biasing high activations and allows the classifier to process data that is normalized across power values. To calculate the compressed power, the power of the frequency bands is first calculated which is shown by Equation III.13.

\[ P_f = \frac{1}{N} \sum_{n=1}^{N} (x_{n_f})^2 \]  

(III.13)

where \( x_{n_f} \) is the magnitude of the \( f \)th frequency band of the \( n \)th sample, \( N \) is the number of samples, and \( P_f \) is the power in a given frequency band. Compressing the estimated power values by using the logarithm of the power is done as shown in Equation III.14.

\[ LP_f = \log(P_f) \]  

(III.14)

Where \( LP_f \) is the logarithmic power of the signal in a given frequency band. With the type of method used for Recoursing Energy Efficiency, applying the normalization techniques to the logarithmic power can be done to get two other features: the normalized logarithmic power (NLP) given by Equation III.15 and the absolute logarithmic normalized logarithmic power (ALNLP) given by Equation III.16.

\[ NLP_f = \frac{LP_f}{LP_{total}} \]  

(III.15)

\[ ALNLP_f = \log\left(\frac{LP_f}{LP_{total}}\right) \]  

(III.16)

III.4 CLASSIFICATION

As described in the previous chapter, it is found that the Multilayer Perceptron Network would offer the best ability to classify the extracted features due to its ability to generalize data and form nonlinear manifolds, which is better than using other methods like K-Nearest Neighbor
(KNN) and Linear Discriminant Analysis (LDA). When using linear methods like LDA, a linear boundary is created to separate different types of data, but since the data may form a nonlinear plane, a linear boundary may be able to classify most of the data, but will not be able to fully classify the data. It is best to have a classifier that forms a nonlinear boundary, especially data that has high variations. Figure III.5 shows how a nonlinear boundary correctly divides the data into two correctly classified datasets. KNN is used to create that nonlinear boundary by finding points that are close to it, but does not have the ability to classify some neighbors correctly. This can be visualized by outliers of different classes being grouped together due to being close. This works with datasets that have good separation, but may not work with smaller boundaries. Therefore, it is decided to use the multilayer perceptron due to the ability to form nonlinear boundaries and group classes correctly when training.
Figure III.5: A nonlinear boundary created to separate a set of data. When a linear boundary is formed, some data may be classified incorrectly.

For the experiments, a multilayer perceptron network is developed with five outputs for each of the five distinct emotions (joy, sadness, disgust, fear, and neutral). Since there are 256-channels of EEG data and separate each channel into 5 different frequency bands (delta, theta, alpha, beta, and gamma), the design of the MLP network contains 1280 inputs. For the amount of hidden layers, it is found that using 20 hidden nodes in the hidden layer gave a good amount of convergence while limiting the ability of the perceptron network to overfit the data.
CHAPTER IV
RESULTS AND DISCUSSIONS

The 256-channel EEG data was collected from five different subjects, which contain five instances for each of the five emotions. For each instance of emotion, the architecture preprocesses the data, performs DWT decomposition, and then segments the data into one-second time intervals. The features are then calculated in each interval, resulting in 750 epochs for each subject. For comparison, several other features are calculated to test the effectiveness of the Recoursing Energy Efficiency algorithm. A two-layer MLP classifier is used with 1280 input nodes for the feature data, 20 hidden nodes, and 5 output nodes corresponding to each of the five emotions. To train the MLP classifier, the feature data is separated into a training set (40% of the data), a validation set (20% of the data), and a testing set (40% of the data) for cross-validation of data. The outputs from the MLP network provided a confidence value, so the highest value among the outputs would be the resulted emotion.

The results are divided up into two different experiments. For our first experiment, the classifier was trained for 10,000 iterations. This allows us to see the classification ability of each of the features presented to the classifier. The number of iterations was chosen to be 10000 for the network because it gave near complete convergence of the network without specifying a lowest error goal, especially since some features may not converge to the lowest error goal. The second experiment allows the network to converge to have a mean-squared error (MSE) of 0.01.
This allows each of the features that can be used to significantly classify emotion to train completely under the constraint of having a lowest error goal.

IV.1 EXPERIMENT 1

Table IV.1 shows the average recognition rate for all of the spectral descriptors and for each emotion. The power descriptor gave the lowest average recognition rate of 47.07%, along with energy and mean spectral descriptor. Since these three descriptors describe the average of the signal, they will provide less ability to see what the difference is across the signals. The root-mean squared (RMS) value gives better recognition while providing an average of the signals due to the slight compression it provides. The next descriptor to give better results is standard deviation, which gives a feature to describe the range of changes in the signal. Since it is best for the descriptors to find variation changes for emotion, these descriptors give better recognition than averaging types, even though this descriptor does still give an average of deviation. The next best features are the maximum and minimum values, giving 72.04% and 75.51% overall recognition. This is surprising because one would think that averaging statistics would give a better indicator of variation changes, but the maximum and minimum of the signals give even better recognition. This may be due to the maximum and minimum values giving an even more compressed range of signals for the classifier to process.

The logarithmic power value has the best overall recognition rate across all emotions, reaching a recognition rate of 91.82%, while having the best recognition rates for four out of the five emotions. The REE, ALREE, NRMS, and ALRMS features provide high recognition rates as well, reaching 83.47%, 90.00%, 86.67%, and 89.40% overall recognition rates across all emotions respectively. This means that using features that normalize and compress the data work
to improve noticeable differences in features while properly quantifying these features for classification.

<table>
<thead>
<tr>
<th>Feature Type</th>
<th>Joy</th>
<th>Disgust</th>
<th>Sadness</th>
<th>Fear</th>
<th>Neutral</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>81.02</td>
<td>74.94</td>
<td>68.40</td>
<td>69.57</td>
<td>83.26</td>
<td>75.51</td>
</tr>
<tr>
<td>Maximum</td>
<td>82.27</td>
<td>65.86</td>
<td>72.05</td>
<td>61.39</td>
<td>77.61</td>
<td>72.04</td>
</tr>
<tr>
<td>Mean [40]</td>
<td>55.61</td>
<td>50.11</td>
<td>71.40</td>
<td>60.18</td>
<td>54.87</td>
<td>58.27</td>
</tr>
<tr>
<td>Standard Deviation [40]</td>
<td>74.94</td>
<td>50.34</td>
<td>42.24</td>
<td>61.90</td>
<td>83.84</td>
<td>62.98</td>
</tr>
<tr>
<td>RMS [44]</td>
<td>78.68</td>
<td>58.72</td>
<td>52.76</td>
<td>50.82</td>
<td>74.54</td>
<td>62.76</td>
</tr>
<tr>
<td>Power [41]</td>
<td>41.52</td>
<td>50.96</td>
<td>40.55</td>
<td>49.46</td>
<td>52.5</td>
<td>47.07</td>
</tr>
<tr>
<td>Energy [44]</td>
<td>63.09</td>
<td>51.28</td>
<td>52.55</td>
<td>54.55</td>
<td>68.34</td>
<td>57.87</td>
</tr>
<tr>
<td>REE [44 – 45]</td>
<td>84.99</td>
<td>91.56</td>
<td>70.80</td>
<td>86.73</td>
<td>83.30</td>
<td>83.47</td>
</tr>
<tr>
<td>NRMS [47]</td>
<td>86.71</td>
<td>89.12</td>
<td>82.68</td>
<td>84.03</td>
<td>91.43</td>
<td>86.67</td>
</tr>
<tr>
<td>ALREE [45]</td>
<td>89.72</td>
<td>90.40</td>
<td>88.67</td>
<td>90.46</td>
<td>90.72</td>
<td>90.00</td>
</tr>
<tr>
<td>ALRMS [47]</td>
<td>90.00</td>
<td><strong>93.93</strong></td>
<td>87.82</td>
<td>85.25</td>
<td>90.44</td>
<td>89.40</td>
</tr>
<tr>
<td><strong>LP</strong></td>
<td><strong>91.83</strong></td>
<td>91.78</td>
<td><strong>90.49</strong></td>
<td><strong>90.71</strong></td>
<td><strong>94.36</strong></td>
<td><strong>91.82</strong></td>
</tr>
</tbody>
</table>

Table IV.1: Recognition rates for classical, normalized, and logarithmic power feature metrics for Experiment 1. Logarithmic power gave the highest recognition of 91.82% across all emotions

**IV.2 EXPERIMENT 2**

Table IV.2 shows results for normalized, logarithmic, and hybrid feature metrics. The hybrid feature metrics, normalized logarithmic power (NLP) and the absolute logarithmic normalized logarithmic power (ALNLP), offer a way to combine both the compression ability of the logarithmic power algorithm and the normalization of the REE algorithm.
<table>
<thead>
<tr>
<th>Feature Type</th>
<th>Joy</th>
<th>Disgust</th>
<th>Sadness</th>
<th>Fear</th>
<th>Neutral</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>REE [44 – 45]</td>
<td>89.32</td>
<td>93.62</td>
<td>81.56</td>
<td>92.07</td>
<td>88.63</td>
<td>88.87</td>
</tr>
<tr>
<td>NRMS [47]</td>
<td>92.21</td>
<td>92.38</td>
<td>88.81</td>
<td>91.13</td>
<td>89.32</td>
<td>90.80</td>
</tr>
<tr>
<td>ALREE [45]</td>
<td>95.70</td>
<td>89.23</td>
<td>89.21</td>
<td>89.97</td>
<td>89.03</td>
<td>90.53</td>
</tr>
<tr>
<td>ALRMS [47]</td>
<td>94.74</td>
<td>92.62</td>
<td><strong>91.13</strong></td>
<td>90.61</td>
<td>89.58</td>
<td>91.73</td>
</tr>
<tr>
<td>LP</td>
<td><strong>96.25</strong></td>
<td>94.70</td>
<td>90.72</td>
<td><strong>94.88</strong></td>
<td><strong>94.57</strong></td>
<td><strong>94.27</strong></td>
</tr>
<tr>
<td>NLP</td>
<td>93.79</td>
<td><strong>95.11</strong></td>
<td>87.00</td>
<td>91.46</td>
<td>93.73</td>
<td>92.20</td>
</tr>
<tr>
<td>ALNLP</td>
<td>90.03</td>
<td>92.56</td>
<td>90.78</td>
<td>88.59</td>
<td>87.38</td>
<td>89.87</td>
</tr>
</tbody>
</table>

Table IV.2: Normalized and logarithmic power feature metric recognition rates for Experiment 2. The logarithmic power feature metric provided the highest average recognition rate of 94.27% across all emotions.

From Table IV.2, noticeable improvement is found in recognition capabilities derived from the Recoursing Energy Efficiency algorithm. From the two variations from REE, the normalized RMS feature and the normalized logarithmic power feature, there is much improvement due to the compression of the features before normalization. Since logarithmic power gives more compression than RMS and gives even higher recognition rates, higher compression gives better recognition rates for spectral features. Noticeable improvement is found in recognition from features derived from ALREE. The ALRMS algorithm gives even better results than the ALREE algorithm, which may mean further compression would provide better recognition, but since the ALNLP gives even lower recognition results, there has to be a compromise between the normalization of the data and the compression of the data.

The REE algorithm gives normalization across the features in different frequency bands and the ALREE gives compression of the features after the normalization. When we apply compression before the normalization, higher recognition rates is obtained but since compression is applied, normalization, and then compression to a couple of the features, over-compression is
done on the features, thus compromising the recognition ability of the feature. Instead of using combinations of compression and normalization, utilizing full compression of features gives even better recognition than other feature types.

From Table IV.1 and Table IV.2, logarithmic power results in 2% higher recognition than all other algorithms. Using the hybrid techniques, NLP and ALNLP, adding normalization to the compressed logarithmic power feature hinders the recognition rate by almost 2%, but still provides higher rates than other methods. Therefore, using logarithmic power as a compressed feature is better than other compression and normalization features.

IV.3 HISTOGRAM ANALYSIS

Figure IV.1 shows the histograms of the RMS, ALRMS, and the LP feature metrics for all of the different frequency bands. The multilayer perceptron network usually normalizes the dataset used for training and testing. Since it takes the maximum and minimum values of the datasets and normalizes the dataset to values between 0 and 1, values must be entirely distributed throughout the whole range of the classifier which aids in the classification of the data. The normalization and compression techniques transform the data to be entirely distributed into the form of a Gaussian distribution. The normalization uses the combination of all the features given to provide a better distribution, while the logarithmic compression technique compresses values that are high in magnitude while stretching values that are smaller. This type of technique is seen in image enhancement, where nonlinear gamma correction is used to bring out features in both underexposed and overexposed images.
Figure IV.1: Histogram distributions of RMS (left column), ALRMS (middle column), and LP (right column) for different frequency bands.
IV.4 COEFFICIENT OF VARIATION

To quantify the level of dispersion of a distribution, we can calculate the coefficient of variation of the distribution, as shown in Equation IV.1.

\[ c_v = \frac{\sigma}{\mu} \]  

(IV.1)

where \( \sigma \) is the standard deviation, \( \mu \) is the mean, and \( c_v \) is the coefficient of variation of the distribution. Table IV.3 shows the coefficient of variation across all frequency bands for each feature type. The normalization and compression metrics provided coefficients of variation which are less than 1, which results in higher recognition rates. Note that having the lowest coefficient of variation does not give the highest recognition rate, but rather the provided features along with the reasonable distribution of data.

<table>
<thead>
<tr>
<th>Freq. Bands</th>
<th>Power</th>
<th>Energy</th>
<th>RMS</th>
<th>REE</th>
<th>NRMS</th>
<th>ALREE</th>
<th>ALRMS</th>
<th>LP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma (D2)</td>
<td>7.84</td>
<td>7.30</td>
<td>1.17</td>
<td>1.43</td>
<td>0.59</td>
<td>0.29</td>
<td>0.16</td>
<td>0.21</td>
</tr>
<tr>
<td>Beta (D3)</td>
<td>13.50</td>
<td>7.59</td>
<td>1.01</td>
<td>1.12</td>
<td>0.49</td>
<td>0.34</td>
<td>0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>Alpha (D4)</td>
<td>14.02</td>
<td>12.26</td>
<td>0.99</td>
<td>1.00</td>
<td>0.42</td>
<td>0.35</td>
<td>0.20</td>
<td>0.22</td>
</tr>
<tr>
<td>Theta (D5)</td>
<td>18.40</td>
<td>17.58</td>
<td>1.36</td>
<td>0.91</td>
<td>0.39</td>
<td>0.35</td>
<td>0.24</td>
<td>0.33</td>
</tr>
<tr>
<td>Delta (D5)</td>
<td>25.11</td>
<td>21.06</td>
<td>2.19</td>
<td>0.21</td>
<td>0.21</td>
<td>1.09</td>
<td>0.47</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Table IV.3: The coefficient of variation for the different feature metrics across different frequency bands.

The normalized and compression methods (REE, NRMS, ALREE, ALRMS, and LP) gave the lowest coefficients of variation, which correlates to the higher recognition rates found by using these methods.
IV.5 NONLINEAR RANGE COMPRESSION EFFECTIVENESS

The logarithmic power algorithm is used for compressing the entire range of the signal for a given frequency band. High values within the range of the signal were compressed while low values were stretched to offer a Gaussian distribution of values. From Figure IV.1, we can see that the range compression of LP allows a more definitive distribution than the distribution shown from the RMS and ALRMS techniques. The logarithmic function offered the best nonlinear range compression for the data, as compared to the RMS function, which uses the square root of the value. Using range compression helps preserve the relationship of values for all frequency bands while forming a normalized distribution suitable for classification. The normalization technique used by the REE algorithm normalizes across all of the frequency bands, which loses some information regarding the strength of the values with respect to the frequency band. Even though normalization offered better recognition than conventional techniques, range compression offered the best recognition because of the conservation of relationship of the values across different frequencies.
CHAPTER V
CONCLUSION AND FUTURE WORK

In this paper, it was seen that using nonlinear range compression for conventional features metrics resulted in the highest recognition rates over all types of feature metrics (classical, normalized, and hybrid feature metrics). The logarithmic power (LP) algorithm gives 94.27% recognition across all emotions, which was due to the preservation of relationships in the values. Normalization offered high recognition rates also, but due to the loss of some relationships between the different frequency bands, nonlinear range compression provided better results. Other nonlinear range compression techniques are currently researched that may boost the recognition ability of the system.

For future work, methods are being researched to remove eye blink/movement artifacts in EEG. The most prominent type of artifact removal technique that we have researched is Independent Component Analysis, which is a blind source separation technique and is showcased in our latest Conference paper submission. Independent Component Analysis works the best amongst methods like EOG-based eye artifact removal and other blind source separation techniques like Principle Component Analysis. The following figure shows the original EEG with eye blinks present and the corrected EEG using ICA.
Methods that provide three-dimensional source localization representations for the analysis of brain activity from two-dimensional scalp activation representations are also being looked at. Different source localization methods are being currently researched that provide the
best localization of sources in the brain. These methods include Local Auto-Regressive Average (LAURA) and Low Resolution Brain Electromagnetic Tomography. Improvement on these algorithms to further localize and characterize brain activations is desired.
BIBLIOGRAPHY


detection) with event-related brain potentials,” Psychophysiology, vol. 28, no. 5, pp. 531–


[10] Dirlich, G., Vogl, L., Plaschke, M., Strian, F.: Cardiac field effects on the eeg,

elimination from eeg using energy function based transformation, in Information and


W., Van Huffel, S., A new muscle artifact removal technique to improve the interpretation
of the ictal scalp electroencephalogram. In: Engineering in Medicine and Biology

[14] Ferdjallah, M., Barr, R.: Adaptive digital notch filter design on the unit circle for the
removal of powerline noise from biomedical signals, IEEE Transactions on Biomedical

methods for removing eye movement artifact from the EEG, Electroencephalography and


<http://wakkanew.wordpress.com/tag/plutchiks-wheel-of-emotions/>
