Theses and Dissertations   Graduate School

2021

Estimating Affective States in Virtual Reality Environments using the Electroencephalogram

Meghan R. Kumar
Virginia Commonwealth University

ESTIMATING AFFECTIVE STATES IN VIRTUAL REALITY

ENVIRONMENTS USING THE ELECTROENCEPHALOGRAM

A thesis submitted in partial fulfillment of the requirements for the degree of Master

of Science at Virginia Commonwealth University.

by

MEGHAN KUMAR

August 2019-August 2021

Director: Dr. Dean Krusienski,

Professor, Department of Biomedical Engineering

Virginia Commonwewalth University

Richmond, Virginia

August, 2021

Acknowledgements

I would like to thank my advisor Dr. Dean Krusienski for his guidance and support throughout all aspects of the thesis and Master's program. His knowledge and plentiful experience encouraged me in the time of my academic research and daily life. I would also like to thank the other memebers of my committee, Dr. Paul Wetzel and Dr. Yusuke Yamani for their valuable time and constructive comments in regard to the thesis. I would also like to acknowledge my lab group, Connor Delaney, Pedram ZS, and Gabe Popescu for their support throughout the thesis. Finally, I am most grateful for my friends and family especially my parents, for their continued support throughout my entire life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

ESTIMATING AFFECTIVE STATES IN VIRTUAL REALITY

ENVIRONMENTS USING THE ELECTROENCEPHALOGRAM

By Meghan Kumar

A thesis submitted in partial fulfillment of the requirements for the degree of Master

of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2021.

Director: Dr. Dean Krusienski,

Professor, Department of Biomedical Engineering

Recent interest in high-performance virtual reality (VR) headsets has motivated research efforts to increase the user's sense of immersion via feedback of physiological measures. This work presents the use of electroencephalographic (EEG) measurements during observation of immersive VR videos to estimate the user's affective state. The EEG of 30 participants were recorded as each passively viewed a series of one minute immersive VR video clips and subjectively rated their level of valence, arousal, dominance, and liking. Correlates between EEG spectral bands and the subjective ratings were analyzed to identify statistically significant frequencies and electrode locations across participants. Model feasibility and performance was studied using stepwise regression and binary Support Vector Machine models. The model results indicate that scalp measurements of electrical activity can reliably estimate subjective scores of perceived affective states.

CHAPTER 1

INTRODUCTION

1.1 Introduction

As technology advances, portable electronic devices and computers are becoming primary mediums for communication, entertainment, and education. Research interest in enhancing virtual reality and augmented reality systems has also grown in recent years for commercial, military, and education applications. This has resulted in an increase in human-computer interaction (HCI) and a need to further develop these systems to enhance the user experience. Most contemporary HCI systems are deficient in interpreting human factors and lack emotional intelligence primarily because they are unable to identify human emotional states. Affective state estimation can be used by machines to detect emotional cues occurring during human computer interaction and synthesizing emotional responses [1]. This technique can be beneficial in work environments, but also in education, interactive entertainment, athletics, etc. In particular, video games and virtual reality applications could directly react to the operator's affective state and manipulate the environment in order to adjust to optimize performance or the user experience.

There are three main categories of measures to assess affective state in human subjects: self-assessment measures, performance measures, and physiological measures. While self-assessment and performance measures work well to evaluate each scenario, they do not work well for immediate feedback [2]. In contrast, physiological measures stay in the background, allowing for the application of multiple undefined scenarios to return information about the process of the human body. These phys-

iological measures can be combined with both the self-assessment and performance measures, in order to produce an affective state algorithm. The use of various peripheral measures like pupil size, eye movements, eye blinks, respiration, and central nervous system measures like electrocenphologram (EEG) can provide reliable results for affective state estimation. The implementation of physiological activity as an input to a closed-loop control system for providing feedback to the user is known as a brain-computer interface (BCI) [3]. BCIs can be used for active or passive applications; for the purpose of this project a passive application is implemented for affective state estimation.

1.2 Brain-Computer Interfaces

A brain-computer interface (BCI) is a method of communication based upon neural activity generated by the brain and is independent of its normal output pathways of peripheral nerves and muscles [4]. Figure 1 depicts the four main components of a BCI: signal acquisition, feature extraction, feature translation and device control. The goal of BCI is to provide a new channel of output for the brain that requires voluntary adaptive control by the acquisition of neural activity, which can be performed by invasive or noninvasive techniques. Invasive techniques such as the electrocorticogram (ECoG) provide higher spatial resolution and a better signal to noise ratio (SNR), but include risks associated with surgery. While non-invasive techniques including electroencephalography (EEG) or functional near-infrared spectroscopy (fNIRS) suffer from comparatively low SNR and spatial resolution, they can be performed without any penetration of scalp. Signal acquisition is performed, features are extracted using signal processing algorithms such as temporal and spatial filters or frequency spectrum analysis to decode brain signals. Features are then translated to device commands using classification or modeling algorithms to predict
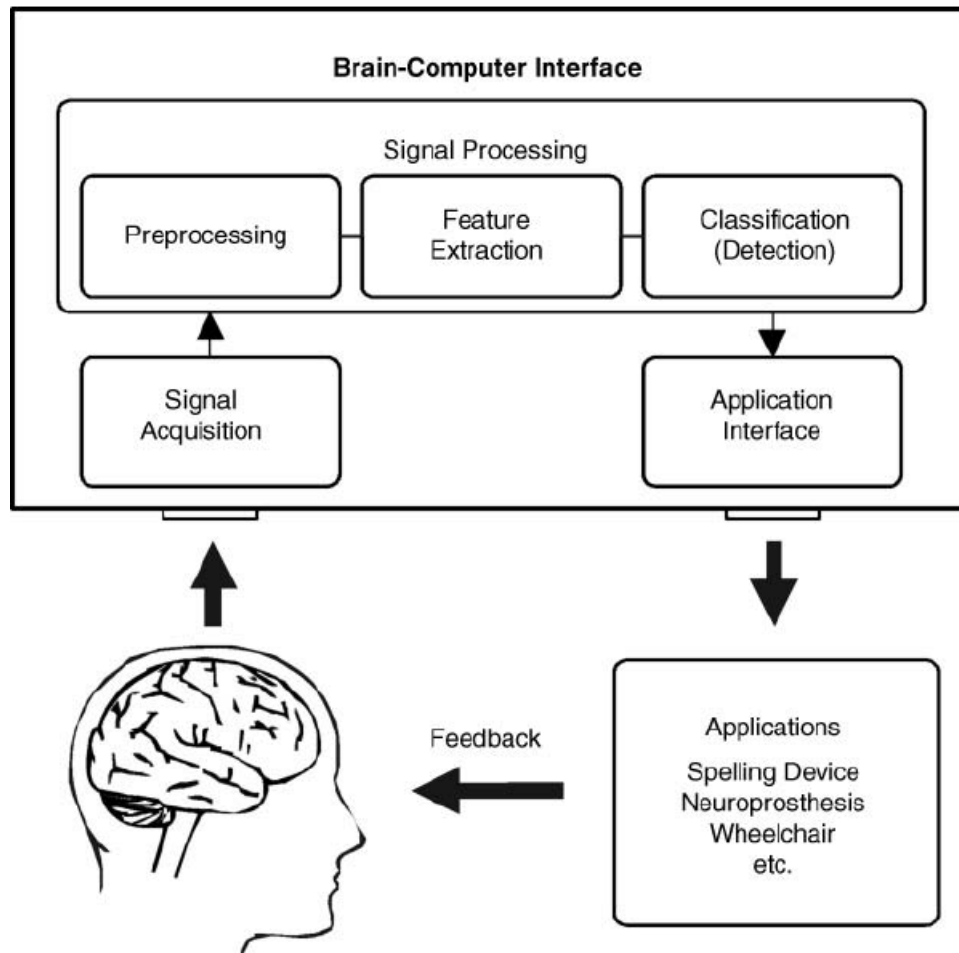
Fig. 1. Components of a BCI system

the user's intentions. These commands serve as the input to external communication and control devices such as robotic arms, wheel chairs, etc. [5].

1.3 Motivation

Room-scale wireless virtual reality (VR) headsets allow users to experience a wide range of realistic and interactive environments. Custom scenarios and tasks have been developed in VR to support various commercial, military, and research applications. While current VR systems provide effective and flexible platforms, fur-

ther development is needed to increase the user's sense of immersion and improve the human-computer interactions. The majority of VR systems use an open-loop structure that does not adapt the environment or task based on human factors other than standard movement tracking and hand controller inputs. Physiological feedback is critical to incorporate factors such as cognitive or affective states into the interactions [6]. This thesis aims to demonstrate that estimates of affective state based on EEG can reliably be obtained while in a immersive VR environment.

1.4 Objective and Approach

The primary objective of this work is to improve the physiological characterization of human affective states using VR and EEG measurements. A study was performed on 30 participants using 13 publicly available 360° VR videos [7]. EEG signals were measured while participants viewed a series of VR videos, each followed by a prompt to complete a Self Assessment Manikin (SAM). The participants provided a rating between 1-9 across four categories in the SAM: valence, arousal, dominance, liking. Correlations between the participant's affective state ratings and EEG frequency bands were computed. EEG electrodes found to be statistically significant in correlating with affective state were identified for future use to support the development of closed-loop VR systems.

The subsequent chapters of this thesis are organized as follows. Chapter 2 presents background of this experiment highlighting emotional state models, the use of electrophysiological activity in affective state analysis. The current state-of-the-art for BCIs in VR environments is also described. Chapter 3 covers background on the four main components of a BCI: signal acquisition, data preprocessing, feature extraction, and classification. Chapter 4 describes the design of the program including the hardware system and experimental design, parameters for data acquisition and

synchronization. Chapter 5 begins the results section with the analysis of affective state ratings. Chapter 6 describes artifact suppression technique applied to control for head movement artifacts. Chapter 7 presents an analysis of the correlations between EEG power bands and SAM ratings. Chapter 8 discusses two affective state estimation techniques using EEG: stepwise regression and support vector machines (SVM). Chapter 9 concludes the dissertation with a discussion of the main contributions and possible future applications of this research.

CHAPTER 2

BACKGROUND

This chapter discusses background on human emotion modeling approaches, the application of BCIs for affective state estimation, and the integration of VR systems with BCIs.

2.1 Affective Neuroscience

Affective neuroscience is aimed to elucidate the neural networks underlying the emotional process and their consequences on physiology, cognition, and behavior. In affective neuroscience, the concept of human emotions can be depicted from various constructs such as feelings, mood, and affects [8]. Feelings can be viewed as a personal experience associating itself with emotion. Moods are diffused affective states that generally last longer than emotions and are less intense. Lastly, affect is an encompassing term describing the topics of emotions, feelings, and moods [9].

Emotions can produce different characteristics indicative of human behavior, and affect decision making, perception, human interaction, and human intelligence. Emotions can also affect human health as well as work efficiency. Three major components which influence the psychological behavior of a human are personal experiences, physiological responses, and behavioral responses [10]. To better grasp the kinds of emotions expressed daily, human emotion must be categorized and quantified in order to be used for feedback in a BCI system. Researchers have proposed various categorical and dimensional models in order to classify different emotions and their levels.

6

2.2 Models of Human Emotion

The categorical model approach revolves around the idea of basic emotions that are imprinted in human physiology. Ekman states there are certain characteristics of basic emotions:(1) humans are born with emotions; (2) humans exhibit the same emotions in the same situation; (3) humans express these emotions similarly; and (4) humans show similar physiological patterns when expressing the same emotion [2]. These consist of surprise, anger, happiness, sadness, fear, and disgust. Facial expressions that depict each of these emotions are shown in Figure 2.
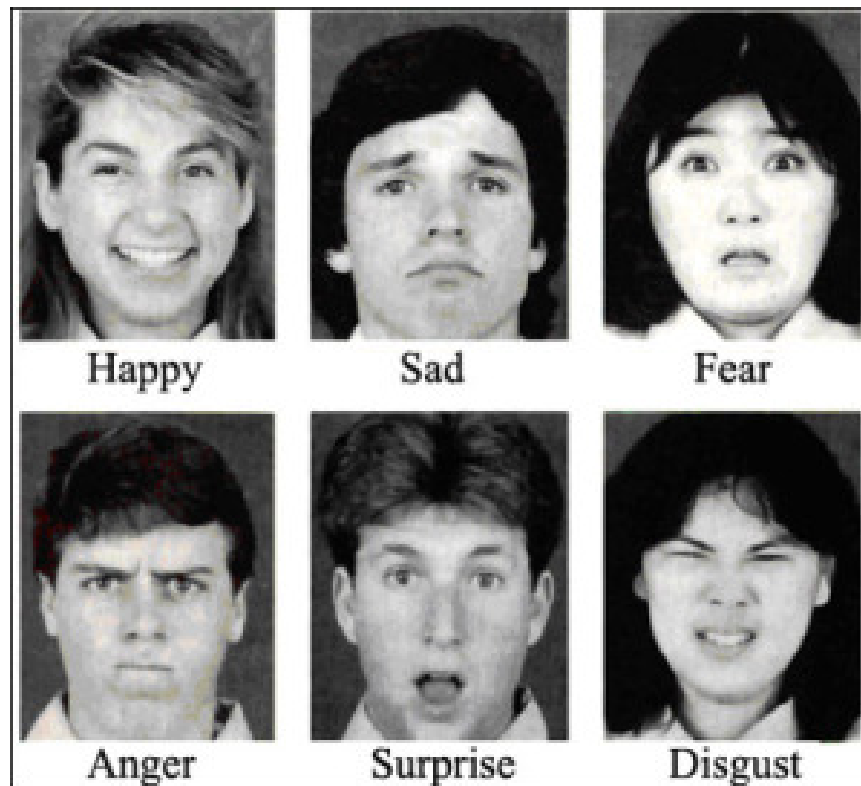
Fig. 2. The six basic emotions by Ekman [2].

These expressions are commonly displayed and recognized facial expressions through out various cultures in the same fashion. Ekman later extended on the list of basic emotions by including amusement, contempt, contentment, embarrass-

ment, excitement, guilt, pride in achievement, relief, satisfaction, sensory pleasure and shame. These emotions acted as mediators amongst the six basic emotions, and provided a deeper understanding of expression in terms of scale.

Plutchik takes a different approach on emotional analysis and divides emotions into primary, secondary, and tertiary levels [11]. He defines 8 basic emotions and depicts them in a graphical representation. Plutchik's wheel representation is formed by 4 couples of bipolar emotions, joy is opposed to sadness, anger to fear, anticipation to surprise and disgust to trust. These emotions are then depicted in a conical shape indicating its relationship between one another. Figure 3 depicts the various levels mentioned in the Plutchik theory .

Dimensional models of emotion include the arousal-valence scale which was first proposed by Russell [12]. The arousal-valence scale is commonly used when researching affective states. Each emotional state can be placed on a two-dimensional plane with arousal and valence as the axes. Arousal can range from inactive to active (bored vs. excited), whereas valence ranges from unpleasant to pleasant. Arousal and valence explain most of the variation between extreme emotional states; a third dimension can arguably be added known as dominance. Dominance ranges from helpless and weak feeling to an empowered influential feeling. Figure 4 represents an illustration of the Russell complex.

In order facilitate self-assessment on dimensional scales, the Self-Assessment Manikin (SAM) is performed. The SAM is a nonverbal pictorial assessment technique that directly measures the various dimensions related to the Russell Complex. For each of the valence, arousal, and dominance dimensions there is a series of manikins visualizing the different values along the axes. For each dimension, participants can select the manikin which most closely expresses their felt emotion. The SAM scale generally ranges from 1-9. Bradley et. al compared reports of affective experience
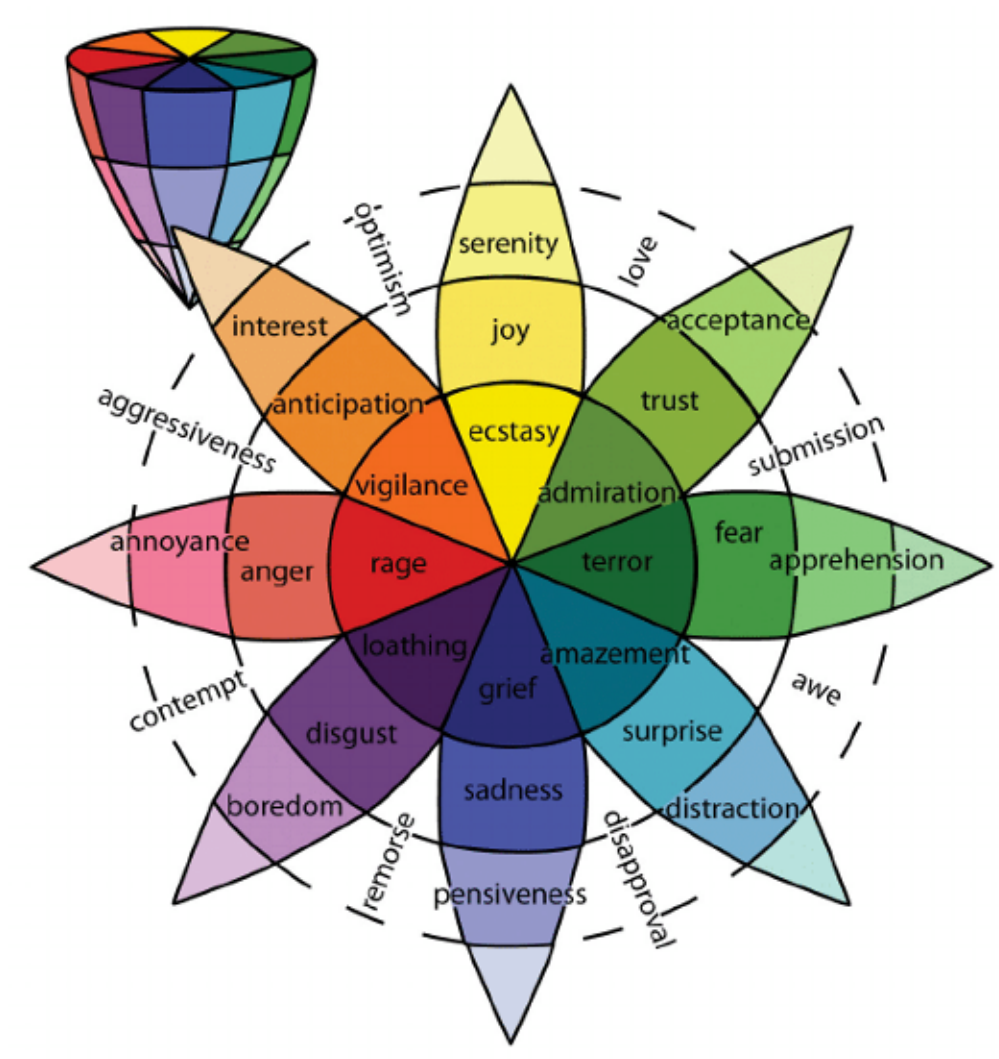
Fig. 3. Plutchik theory of emotion represented in a conical shape [11].

obtained using SAM, requiring three scales: valence, arousal, dominance [13]. Subjective reports were measured to a series of pictures varying in both affective valence and arousal. Correlation across the valence and arousal methods were high however differences were obtained in the dominance dimension, suggesting SAM may better track the personal response to an affective stimulus. It is also an inexpensive method to rapidly assess reports of affective response in many contexts [13]. An example of the pictorial representation of SAM is presented in Figure 5.
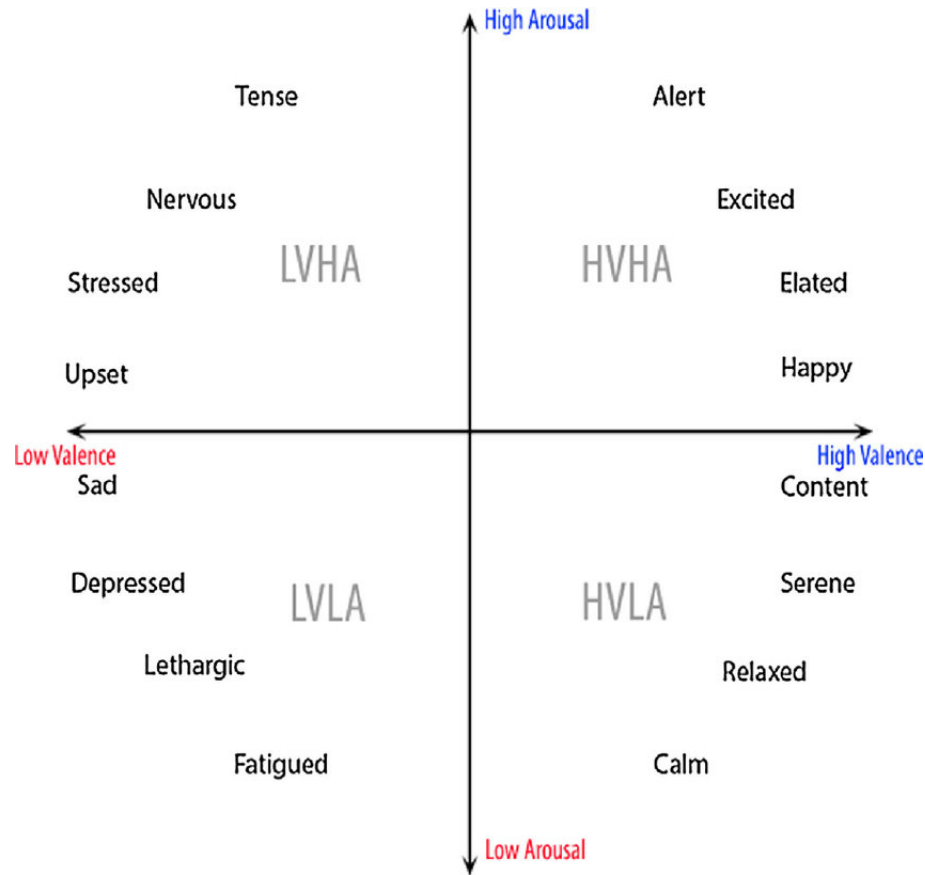
Fig. 4. The Russel Complex [12]

2.3 BCI for Affective State Analysis

BCI for EEG can be classified into two categories, the voluntariness (passive vs. active) and stimulus dependency (independent/induced vs. dependent/evoked) as seen in Table 1. Active BCI is not ideal for affective state estimation because it requires a control interface that would interfere with an affective state task [14]. The main approach for affective state estimation with EEG utilizes passive stimulus-independent BCIs. Passive stimulus-independent affective state estimation uses changes in EEG activity to estimate affective state. Passive stimulus-independent BCIs can run in the background and update without interfering with a task. The technique of

10
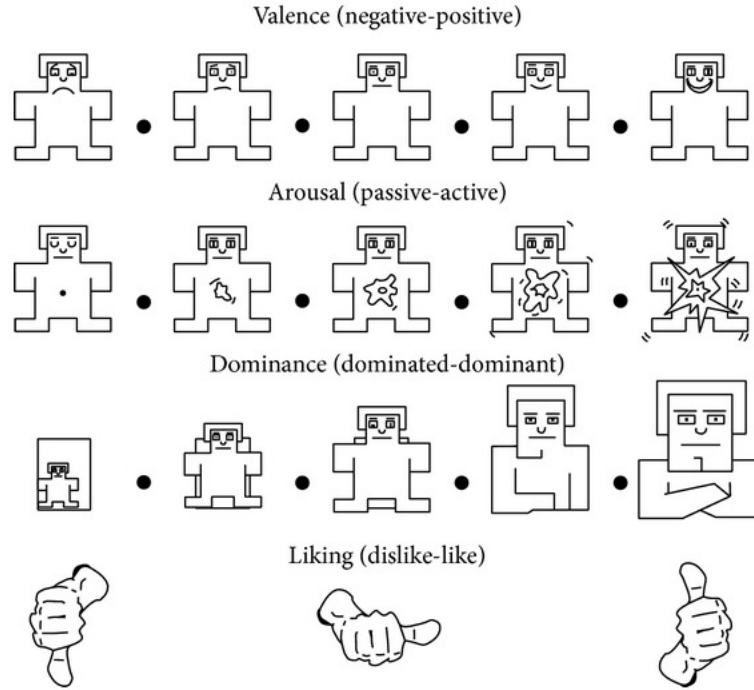
Fig. 5. Self Assessment Manikin(SAM).The first row represents valence, the second row represents arousal, the third row represents dominance [13].

using power spectral bands is commonly used for more continuous experiments [15].

Classification	Stimulus Independent	Stimulus Dependent
Active	Motor Imagery	P300,SSVEP
Passive	Workload Estimation, Emotional Estimation	P300 Workload, Error Related Potentials

Table 1. EEG paradigms classified into voluntariness and stimulus dependency.

EEG compromises a set of signals which may be classified according to their frequency, the most commonly used frequency bands are δ (0-4 Hz), θ (4-7 Hz), α (7-12Hz), β (12-30 Hz), and γ (30-100 Hz) waves.

The delta band (δ) lies below 4 Hz, and is usually only observed within adults in

deep sleep state. When a large amount of delta activity is perceived in awake adults, it can be related to neurological disease [16]. Due to low frequency range, the delta band can easily be confused with artifact signals caused by large muscles of the neck or jaw.

Theta band (θ) ranges between 4-7 Hz, and is normally perceived in awake adults. Large theta frequencies can be seen in young children, and adults in drowsy, meditative, and sleep states. Similar to delta, a large amount of theta activity in awake adults is related to neurological diseases [16]. Theta band is also associated with meditative concentration [17], and a wide range of cognitive processes such as mental calculation, or conscious awareness [18].

Alpha rhythms (α) are primarily found over occipital region in the brain. The alpha band ranges from 8 to 12 Hz, and their amplitude increases when eyes are closed and the body relaxes. The waves attenuate once eyes are opened and mental effort is made. Increasing mental effort causes a suppression of alpha activity, particularly from the frontal areas [19]. The alpha band is also strongly connected to motor activities. These rhythms generally reflect visual processing in the occipital brain region, and can appear to correlate with beta rhythms [20].

The beta band (β) is within 12 to 30 Hz range, and is correlated with motor activity. Beta rhythms are desynchronized during real movement or motor imagery. Beta waves are distinguished by their symmetrical distribution when there is no motor activity present. When movement is present the beta wave attenuates, and the symmetrical distribution changes [21].

The gamma band(γ) spans from 30-100 Hz. The presence of gamma waves in the brain activity of a healthy adult is related to certain motor function or perceptions, among others. Several studies have suggested gamma activity is prominent when presented with visual and auditory stimuli [22, 23]. Gamma rhythms are less commonly

12

used in EEG-based BCI systems, due to artifacts presented such as electromyography (EMG) or electroculography (EOG) affecting the gamma band component of EEG signal [24].

Table 2 displays reported EEG-based functional connectivity in the brain with relationships between specific brain areas and cognitive states. Studies that take single-electrode-level analysis into account have shown that asymmetric activity at the frontal site in the alpha band is associated with emotion. Ekman and Davidson found that enjoyment generated an activation of the brain's left frontal parts [2]. Another study found a left frontal activity reduction when volunteers adopted fear expressions. Increased power in theta band at the frontal mid-line is associated with pleasurable emotions, and the opposite has been observed with unpleasant feelings [15].

Band	Function	Localization
Theta (4-7 Hz)	Associated with subconscious mind and occurs in sleeping and dreaming	Frontal head region and hippocampal region
Alpha (8-13 Hz)	Associated with relaxed mental state or positive emotions yet aware and correlated with brain activation	Parietal and occipital regions
Beta (13-30 Hz)	Associated with active mind and occurs during intense focused mental activity	Motor cortex
Gamma (≥ 30Hz)	Associated with intense brain activity	Various sensory and non-sensory cortical networks

Table 2. EEG frequency bands [25]

.

According to a study involving music video excerpts, it has been observed higher

13

frequency bands such as gamma were detected when subjects listened to unfamiliar songs [26]. Other studies suggest alpha, beta, and gamma frequency bands improve classification of emotion in both valence and arousal dimensions [25]. While averaged results of studies determine frequency ranges for affective state estimation, challenges include finding appropriate patterns amongst all, resulting in the importance of individually tuned classification models [27].

2.4   BCI in VR

In recent years, augmented/virtual reality (AR/VR) technology has advanced, delivering higher levels of immersion and presence to users. AR/VR makes it possible to simulate and evaluate environments under controlled laboratory environments, which can be unfeasible in real space [5]. With the accessibility and flexibility, AR/VR serves as a platform to stimulate "real life" emotional responses. AR/VR technology allows the creation of scenarios testing human emotion that are much more stimulating and expressive than the standard desktop view applications.

In research environments, AR/VR can be used to adjust intensity, complexity, and realism while maintaining full control over experiment. In therapeutic applications AR/VR can create motivating training paradigms which make use of "gamification" approaches. Common passive BCI research approaches tackling adaptive AR/VR technology include attention, workload, and emotion/affective state. Vortmann et al. performed a study on the classification of internal from external attention in AR setting [28]. The research developed a novel AR task requiring continuous spatial alignment, mimicking typical AR interactions. They demonstrated a real time implementation of the attention model allowing online adaption of AR-based user interfaces, such as smart home control in AR using steady state visually evoked potentials. [5, 28].

14

Tremmel et al. performed a study measuring mental workload in VR. For the purpose of the study the classic n-back task was performed, and determined workload levels could be discriminated from the scalp recording despite large level of physical movement [14].

To track the emotional responses wearable EEG can be combined to record physiological signals and then evaluate the mental/affective state. Moreover, the ability of VR to induce emotions has been analysed in studies which demonstrate that virtual environments do evoke emotions in the user [29]. From a study conducted by Malandrakis et al., VR application was implemented to track moods. Based upon their emotional levels while viewing movies, a list of databases for movie recommendations was outputted [30]. Recent papers collected from 2017-2019 found that the common approach towards stimulating user's emotional experience was music, music videos, pictures, video clips, and VR. Of the five stimuli virtual reality has the highest common usage for emotional classification followed by music, music videos, and video clips [25].

CHAPTER 3

SIGNAL PROCESSING FOR AFFECTIVE STATE ANALYSIS

This chapter provides details about the signal acquisition process and commonly used pre-processing and feature extraction techniques for affective state analysis. Additionally, various feature translation techniques are discussed along with model structure identification techniques. The chapter concludes with relevant statistical metrics for analysis. These concepts will be applied in following chapters.

3.1   Signal Acquisition

The EEG signal represents a voltage measured between electrodes placed on the objective region of the scalp and a reference electrode, with respect to a ground electrode. The reference and ground electrodes are generally placed on locations of the scalp (mastoid, earlobe, forehead etc.) that do not induce additional interference onto the desired signal. When collecting EEG data, the sampling rate must always be at least twice as high as the expected frequency being observed in order to satisfy the Nyquist frequency criterion. EEG amplifiers generally record EEG signals from 125-512 Hz to capture the relevant frequencies and avoid aliasing [25].

In order generate reproducable EEG recordings, the standard electrode 10-20 international system is used [31]. These electrode spacings are represented by either 10% or 20% of the total right-left or front-back distance of the skull. The letters of the electrodes represent the region of the brain they are placed on ("F" for frontal lobe, "O" for occipital lobe, etc.). The electrodes for affective state analysis are usually placed on the frontal parietal areas due to the correlation of emotional activity in the

amygdala and frontal lobe. Studies have shown that the frontal scalp exhibits more emotional activation compared to other regions of the brain such as temporal and occipital [32].

3.2 Data Pre-processing

The pre-processing of EEG signals primarily involves signal cleaning and enhancement. EEG signals are often contaminated with noise from internal and external sources. Therefore common pre-processing techniques are essential to suppress noise contamination that could affect EEG interpretation and decoding. The body produces electrical impulses through blinking, muscular movements, and heartbeat that can mix with EEG signals. Artifacts present in the EEG signal must carefully be eliminated or suppressed to ensure affective state information is not contaminated. If spatial or temporal filters are applied, it is imperative not to induce additional signal distortion. Common pre-processing filters used in EEG include bandpass and notch filters. EEG bandpass filters generally filter frequencies between 0.1 and 30-60 Hz. Notch filters are used to attenuate a specific frequency rather than a range. For EEG, the filter eliminates frequencies associated with power line noise (50 to 60 Hz, depending on the specific country) [33].

Since EEG can represent the activity of many different neural populations that might create noise for a task, dimensionality reduction and source separation algorithms are used. Common techniques include principle component analysis (PCA) [34, 35] and independent component analysis (ICA) [36]. Various sub-populations of neurons in the brain can activate synchronously and independently. Additionally, scalp tissue acts as a volume conductor for electrical activity, further mixing the various sources of brain activity and other electrophysiological activity at any given instant [37].

PCA is a statistical feature extraction method that uses linear transformation to convert a set of potentially correlated observations into a set of uncorrelated variables called principal components. The linear transformation generates a set of components from the input data, sorted according to their variance in such a way that the first principal component represents the highest variance, which can represent relevant aspects of the signal or noise sources. PCA is often combined with ICA, as the representation of the data by principle components simplifies and accelerates the application of ICA algorithms [38].

ICA attempts to recover desired signals as a linear mixture of independent source signals, which is referred to as blind source separation (BSS). ICA has traditionally been used as a pre-processing tool before feature extraction in order to remove artifacts present data [39]. The ICA algorithm derives independent sources from highly correlated EEG signals statistically and does not regard to the physical location or configuration of the sources generators of EEG signals. Independence is obtained through minimization of mutual information or maximization of non-Gaussianity [40]. Figure 6 depicts the effects of ICA algorithm on sample EEG data. The original EEG signal contains large amplitude spikes that are recognized as artifacts. After performing ICA, these spikes are effectively suppressed. The left spectral power plot in Figure 6 also depicts the effects of performing ICA on the power spectrum. The power of the lower frequencies due to the artifacts is decreased, while preserving the EEG alpha peak [41].

3.3 Feature Extraction

Features are generally distinguished in the spatial, time, frequency, and/or time-frequency domains. While the use of time-domain features of EEG is not predominant for emotional state detection, multiple approaches exist to determine characteristics
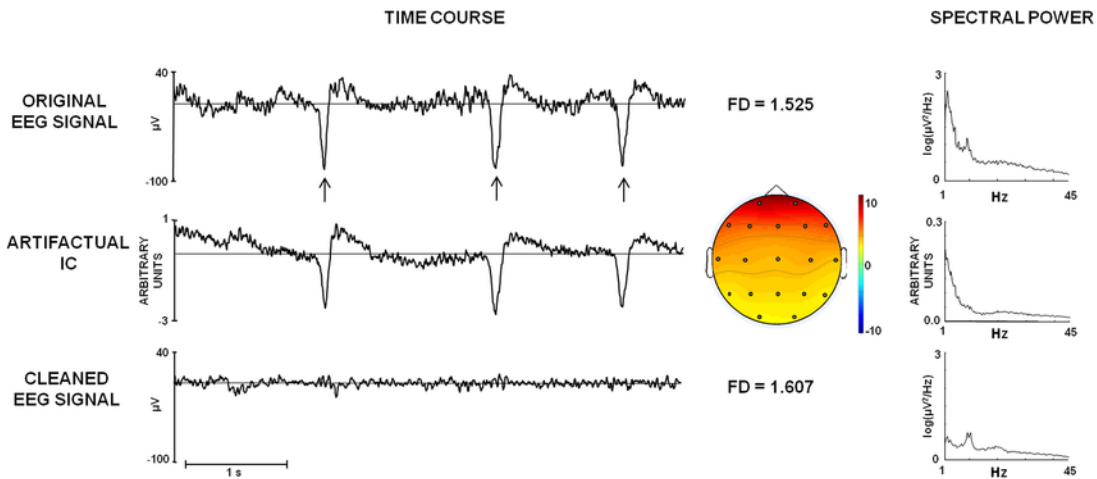
Fig. 6. Example of original EEG signal, an artifact independent component, and cleaned EEG signal. The right panel shows the associated spectral power [41].

of time series that vary between different emotional states. Event-related-potential (ERP) measures the synchronous activity of large populations of neurons in response to a specific event. In a typical ERP experiment, stimuli of various types will be repeated many times.The ERP is computed by averaging the stimulus-locked EEG responses across the same stimulus or category of stimuli to improve the signal-to-noise ratio. Lithari et al. used amplitude and latency of event related potentials (P100, N100, N200, P200, P300) as features in their study implementing pictures to evoke emotion [42]. In an online application, however, it is difficult to detect ERPs related to emotions since the onset is usually unknown (asynchronous BCI). Thus, frequency domain features including band power analysis is the most popular in context of emotion recognition.

The selected frequency bands of EEG vary slightly amongst studies, with the bands commonly defined as given in Table 2. The frequency bands are often computed using fast Fourier transform (FFT) or variations such as short-time Fourier transform (STFT) or the estimation of power spectra density (PSD) via Welch's method. PSD

identifies the power distribution of a signal over the range of frequencies. PSD estimates are defined by parametric or non-parametric approaches. Welch's method improves the accuracy of the classic periodogram, by smoothing over non-systematic noise and is robust to non-stationary signals. In Welch's PSD method, data segments can overlap, and contain windows. The window length has significant impact on the classification as longer windows produce better classification results [27]. Consider a discrete time signal s with N samples:

$$s = x[1], x[2], ...x[N] \tag{3.1}$$

The signal can be separated in K smaller intervals with length M and overlap V:

$$1 : s_1 = x[1], x[2], ..., x[M] \tag{3.2}$$

$$2 : s_1 = x[M - V + 1], x[M - V + 2], ..., x[2M - V] \tag{3.3}$$

$$K : s_k = x[(K - 1)M - (K - 1)V + 1], ..., x[KM - (K - 1)V] \tag{3.4}$$

where $s_i = s_i[1], ..[2], ...s_i[M]$ represents the ith window and K is number of the involved window in PSD calculation. The DFT is calculated for each window

$$S_i[v] = \sum_{m=1}^{M} s_1[m] * w[m] exp\left(\frac{-2\pi j m v}{N_f}\right) 1 \leq v \leq N_f \tag{3.5}$$

where $w = w[1], w[2], ..w[M]$ is the windowing vector, N_f is the DFT size, and $S_i = S_i[[1], S_i[2], ..S_i[N_f]$ represents the vector of frequency samples of ith input window. Periodogram values are squared of the absolute value of the DFT samples:

$$P_i[v] = 1/C \ |S_i[v]|^2 1 \leq v \leq N_f \tag{3.6}$$

where C is normalization factor:

$$C = \sum_{m=1}^{M} w^2[m] \tag{3.7}$$

Calculated periodogram values from different windows are averaged and the Power Spectral Density estimate is obtained:

$$PSD = \left(\frac{1}{K}\right) \sum_{i=1}^{K} P_i[v]1 \leq v \leq N_f \tag{3.8}$$

The average order of K is controlling the dependency on the past. Larger K values allow the system to estimate each frequency component based on more observations. Fast variation of a frequency component leads to greater averaging, which can be a drawback for estimation. Therefore, a robust monitoring system needs to include the variable K to account for the temporal dynamics. When targeting a low frequency under noisy conditions, a K with a larger value can benefit the detection [43].

3.4   Statistical Metrics

Statistical measures such as Pearson's correlation and Fisher's method are commonly used when performing affective state analysis. Correlation is a bi-variate analysis that measures the strength between two variables. In correlated data, the change in the magnitude of one variable is associated with a change in the magnitude of another variable, either in the same (positive correlation) or in the opposite (negative correlation) direction. Pearson's correlation coefficient measures dependence between two variables. Equation (3.9) presents the formula for Pearson's correlation coefficient:

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})}{(n-1)S_x S_y} \tag{3.9}$$

where $\bar{x}$ represents the sample mean for the first variable, S_x is the standard deviation for the first variable, $\bar{y}$ represents the sample mean for the second variable, S_y is the standard deviation of second variable following the n which is the column length [44].

The value of the correlation coefficient lies between -1 and 1. If no correlation exists the value is zero. Because Pearson's correlation assumes normal distributions, other types of correlations can also be computed for non-normal data: Kendall rank correlation, Spearman correlation, and Point-Biserial correlation [45]. Spearman's correlation coefficient is the ranked form of Pearson's correlation describing a monotonic relationship. Martinez et al. performed statistical analysis using Spearman's correlation between EEG features and emotional information present in video games [46]. The study concluded EEG features have strong correlation within arousal and valence scores resulting in a large correlation coefficient. The correlation coefficient is often combined with the p-value when the correlation coefficient is zero (null hypothesis). A small p-value suggests that the null hypothesis is false. If the probability is lower than the conventional 5% ($p \leq 0.05$) the correlation coefficient is called statistically significant.

The formula for Pearson's correlation coefficient uses the t-distribution. Equation 3.10 depicts the formula for p-value, where r represents correlation coefficient and n depicts the number of observations. The p value is 2 X P(T>t) where T follows a t distribution with n-2 degrees of freedom [47].

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \tag{3.10}$$

Fisher's method, also known as Fisher's combined probability test, is a technique for data fusion or meta analysis. Fisher's method combines extreme value probabilities from each test (p-values), into one test statistic (X^2) using the formula in equation

3.11.

$$X_2^2 \ -2\sum_{i=1}^{k} ln(p_i) \tag{3.11}$$

where p_i is the p-value for the ith hypothesis test. When the p-values tend to be small, the test statistic (X^2) will be large suggesting that the null hypothesis are not true for ever test [48].

3.5 Feature Translation

The aim of feature translation in a BCI system is recognition of a user's intention on the basis of features that characterize brain activity provided by the feature extraction step. Regression or classification algorithms can be used to achieve this goal. Regression algorithms employ the features extracted from EEG signal's as independent variables to predict user intentions along a continuum. Classification approaches use the features extracted as independent variables to define boundaries between the different targets in feature space [49]. Figure 7 depicts how the classification model aims to model the separation between classes while the regression model aims to find the trend in the data.

Classification algorithms identify the category of new observations based on the category of previous observations. Classification algorithms are generally trained by dividing previously observed data into independent training and testing sets. Commonly, 70% to 90% of the data is used for training and the remainder for testing. In order to better represent variability in the training and testing data, cross-validation is often applied. Cross-validation effectively repeats the classifier evaluation process by using different subsets of training and testing data. The most common cross-validation approach is "k-fold cross validation," which divides the dataset into k

Classification   **Regression**
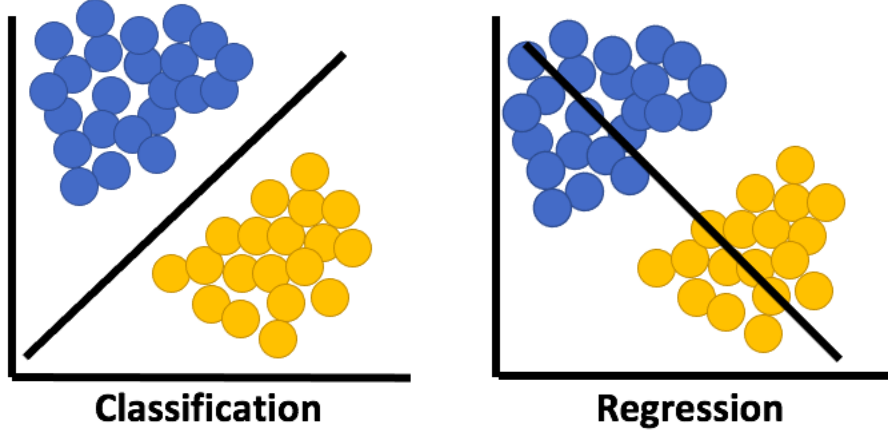
Fig. 7. Classification versus regression modeling approaches.

parts and respectively uses each combination of k-1 parts for training and the remaining part for testing. The k results are then averaged to determine performance of classification method utilized [50].

Commonly used classification algorithms for brain-computer-interface related research include variations of linear discriminant analysis (LDA), support vector machines (SVMs), and artificial neural networks (ANNs). Linear classifiers such as LDA and linear SVMs categorize new observations based on a linear combination of feature values, whereas nonlinear classifiers such as nonlinear SVMs and ANNs rely on a nonlinear combination of features [51].

LDA is a simple classifier that provides acceptable accuracy without high computation requirements. LDA works similarly to PCA in trying to map a dataset into different coordinate system. In contrast to PCA, LDA creates new components that maximize the separability amongst classes. Maximizing the separability can be denoted as:

$$max \left(\frac{(\mu_1 - \mu_2)^2}{\sigma_1^3 + \sigma_2^3} \right) \tag{3.12}$$

With μ as the mean and σ as the variance of the respective classes. LDA uses a

24

hyperplane to separate the classes (multiple hyperplanes can be used for multi-class problems). Data points falling on either side of hyperplane can be attributed to different classes. LDA has been used successfully in numerous BCI systems, such as P300 speller, multiclass, or synchronous BCIs [52]. However, LDA can lead to erroneous classifications in the presence of outliers or strong noise. LDA is usually applied to classify patterns into two classes, but it is possible to extend the method to multiple classes.

The objective of SVM is similar to LDA, but the approach to finding the optimal separating hyperplane is different. In contrast to LDA, SVM selects the hyperplanes that maximize the margins, the distance between the nearest training samples and the hyperplanes [53]. An SVM finds a separating hyperplane with maximal margin. An SVM uses regularization, in order to prevent the classifier from accommodating possibly noisy data sets. SVM has been used to classify features vectors for binary and multiclass problems, and has been successfully used in synchronous BCIs [54].

It is also possible to create a SVM with non-linear decision boundary using a kernel function $K(x,y)$, according to Cover's theorem on the separability of patterns [55]. Cover's theorem states that a complex classification problem cast in a high-dimensional nonlinear space is more likely to be linearly separable than in a low-dimensional nonlinear space. Figure 9 illustrates the concept of an SVM kernel function for which circles and squares denote two-class samples. In order to classify linearly, the kernel function converts the input samples into high-dimensional space. The distance between the dashed separating hyperplane and the dotted lines defined by the support vectors is the margin.

The kernel commonly used in BCI field is Gaussian or Radial Basis Function (RBF) [56]. Non-linear SVM leads to a flexible boundary in data space, increasing classification accuracy. SVM is widely used BCI , because it is simple classifier that

performs well and is robust to the dimensionality of the problem, therefore large training sets are not required.
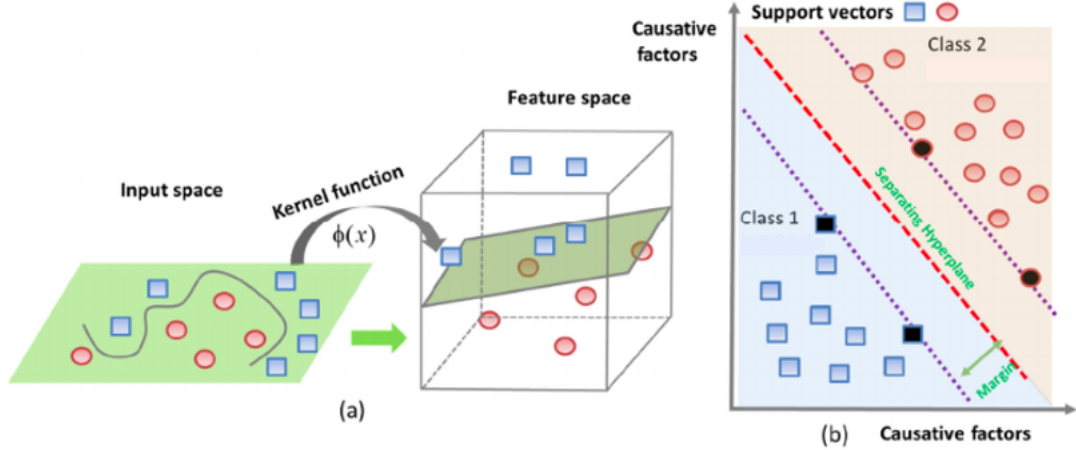
Fig. 8. Support Vector Machine(SVM): (a) Input space is mapped to the feature space with kernel function; (b) Separating hyperplane and margin for classification [57].

In contrast to classification models that decide between a finite set of discrete classes, regression methods are used to estimate an output along a continuum. A common approach to establish regression models is based on least-squares regression. 3.13 depicts the relationship between the response variable, z, the matrix containing column vectors of regressors, X, and the measurement error, v, θ is the estimate model parameters.

$$z = X\theta + v \tag{3.13}$$

The parameter estimates are computed in order to reduce cost function, this results in the solution given in 3.14 for the parameter estimate, $\hat{\theta}$.

$$\hat{\theta} = (X^T X)^{-1} X^T z \tag{3.14}$$

Extending linear regression models for over-determined problems commonly en-

countered in BCI with many possible features and relatively few observations, stepwise regression analysis is able perform joint feature selection and regression [49]. When performing stepwise regression, a set of regressors is defined based on the training data. Stepwise regression uses statistical metrics to generate a robust model by varying terms included in the model. During the stepwise regression routine, a new term is either introduced or a term is eliminated based on its statistical significance. The partial F statistic, F_0, and the partial correlation to the response variable r, are statistical metrics used to analyze the quality of individual model terms. The partial F statistic given in 3.15 describes the importance of the candidate regressor currently included in the model in terms of the parameter estimate and the standard error [58].

$$F_{0_i} = \left(\frac{\theta_i^2}{\sigma_i 2} \right) \tag{3.15}$$

While the partial correlation given in 3.16 determines the significance of each regressor outside of the model structure.

$$r_i = \left(\frac{(x_i - \bar{x})^T (z - \bar{z})}{\sqrt{(x_i - \bar{x})^T (x_i - \bar{x})((z - \bar{z})^T (z - \bar{z})}} \right) \tag{3.16}$$

A metric used to determine the overall fit is the coefficient of determine, R^2, and is given in 3.17 R^2 will increase with an increase in the number of terms included into the model, but the accuracy of the parameter estimates may decrease as the number of regressors included in model increases.

$$R^2 = \left(\frac{y^T z - N\bar{z}^2}{z^T z - N\bar{z}^2} \right) \tag{3.17}$$

This iterative process is followed to select features to be included in the model.

The main difference between using a regression technique versus classification is the output variable in regression is numerical (or continuous) while for classification it is categorical (or discrete). Regression models are commonly evaluated using R-

squared while classification models can be evaluated using a variety of performance metrics such as classification accuracy.

CHAPTER 4

EXPERIMENT DESIGN AND PROTOCOL

This chapter presents details on the screening process for participating in the experiment as well the procedure for the experimental set-up. Details on the videos chosen for the experiment, the VR environment, and equipment used are also described within this chapter. Additionally, the experimental task and data collection process using the HTC VIVE and g.NAUTILIS EEG cap are discussed.

4.1 Participants and Experimental Setup

Thirty healthy individuals (ages 19-34, mean 23.8, 17 females, 13 males) were recruited to participate in the experiment, which was approved by the Institutional Review Board of Virginia Commonwealth University. Each participant completed a screening process consisting of an informed consent form, demographic information form, Motion Sickness Susceptibility Questionnaire (MSSQ) short form [11], and were given instructions describing the Self Assessment Manikin (SAM) affective states (Appendix C). All participants satisfied the criteria by scoring a minimum of 19 on the MSSQ.

Following the screening process, participants were introduced to the HTC VIVE hardware system, which is depicted in Figure 9. The HTC VIVE hardware system consists of a motion tracked headset display, and two "lighthouse" base stations that provide external 6 Degree of Freedom (6DOF) tracking. The VIVE wireless adapter was used in conjunction with the wireless EEG headset such that the participant was untethered and free to rotate to view the VR environment. Figure 10 depicts a

participant wearing the EEG headset and HTC VIVE VR headset.

A visual calibration was performed using the HTC VIVE to correct for lens distance. The wireless 32-channel EEG cap was then placed on the participant's head and the EEG electrodes were filled with electrolyte gel. The electrode cap was covered with a protective plastic hair cap to protect the VR headset from the gel. The VR headset was placed over the EEG cap and the headset was tightened to comfortably fit the participant. Participants were then positioned approximately 1 meter from the recording computer in a seated position, in a swivel office chair. Participants were provided guidelines to complete the SAM in terms of valence, arousal, dominance and liking prior to the start of the experimental task.

Fig. 9. Wireless HTC VIVE headset with two lighthouse base stations [59]

4.2 Experimental Task

The participant's experimental task was to passively observe stimuli within the VR environment and provide subjective affective state ratings using a SAM. In this
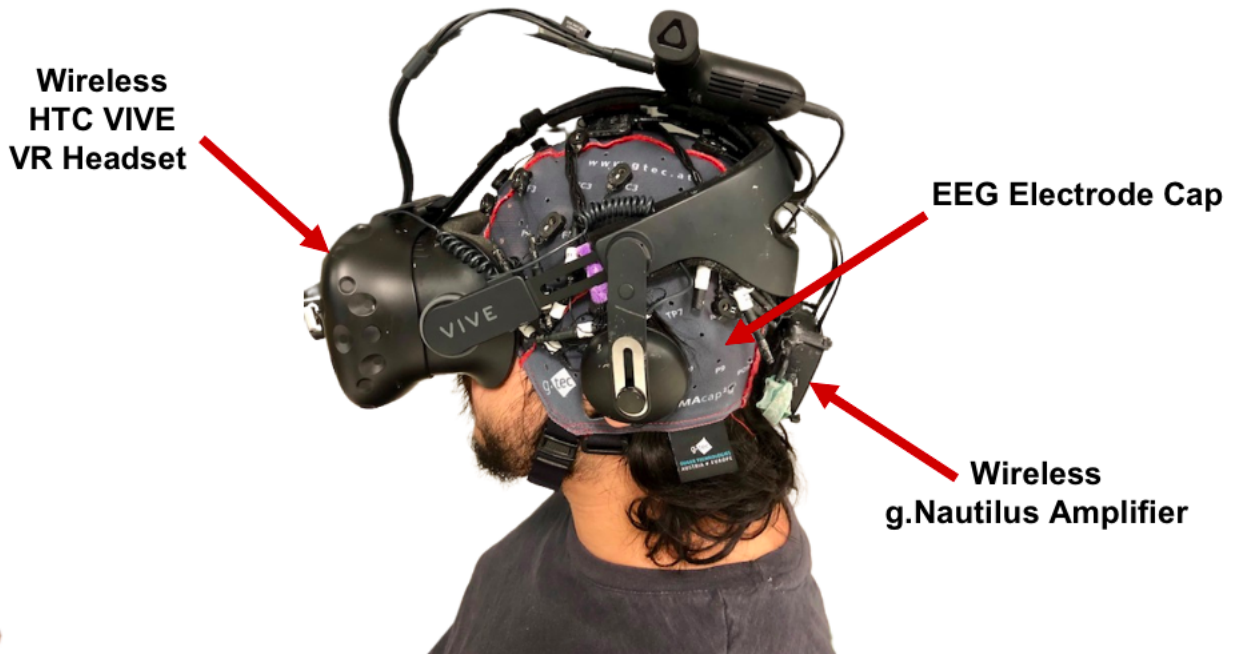
Fig. 10. HTC VIVE and g.Nauilus EEG placed on participant.

experiment, the stimuli are a series of 360° videos viewed through the VR headset to induce various levels of affective states. The videos were obtained from a public database which were used in a previous human emotion study [7]. The videos from the database were organized into three categories based on the subjective ratings from [7]: High-Valence-Low-Arousal (HVLA), High-Valence-High-Arousal (HVHA), and Low-Valence-Low-Arousal (LVLA). Each category consisted of 4 one-minute videos for a total of 12 video segments, and one neutral video to collect baseline EEG data at the beginning of the experiment. The video categories, names and descriptions are provided in Table 3. A visual representation of video stimuli selected and SAM scale in Unity is depicted in Figure 11.

At the start of the experiment, the participant viewed the neutral video and completed the SAM. The remaining 12 videos were then shown in a randomized order for
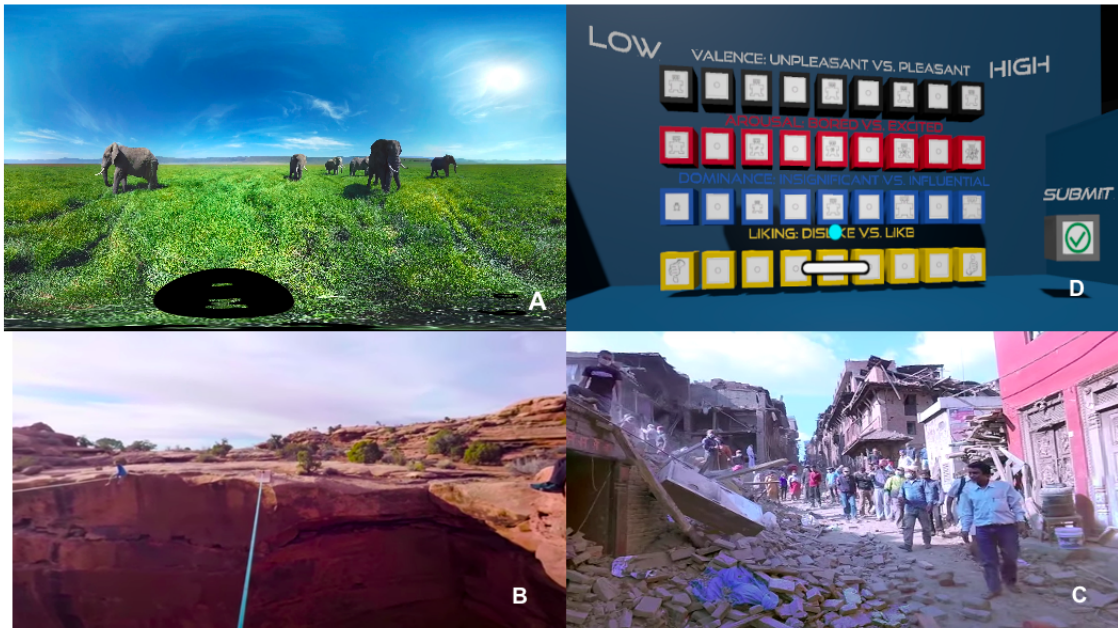
Fig. 11. Snapshots of selected visual stimuli selected and Self Assessment Manikin in Unity. A) Surrounded by Elephants B) Walk the Tightrope C) Nepal Earthquake D) Self Assessment Manikin.

each participant. For a particular trial, the participant's task was to passively view a video, with the ability to rotate the chair and the headset to view the 360° environment from different perspectives. Immediately following the video, participants were prompted to complete the SAM using a scale of one to nine for valence, arousal, dominance, and liking. The icons displayed for the SAM are shown in Figure 5 of Chapter 2. Each row represents a different affective state, and each column represents a different rating. The participant selects the ratings by directing his or her head toward the desired manikin and maintaining the cursor over the manikin for a 5-second dwell time, as indicated by a progress bar. A twenty-second interval between trials was used to collect baseline EEG data. During this interval, participants were asked to keep their eyes open and remain still. Figure 12 provides a diagram of the experimental task. The total duration of the experiment was kept to 25 minutes to

reduce the risk of simulator sickness.
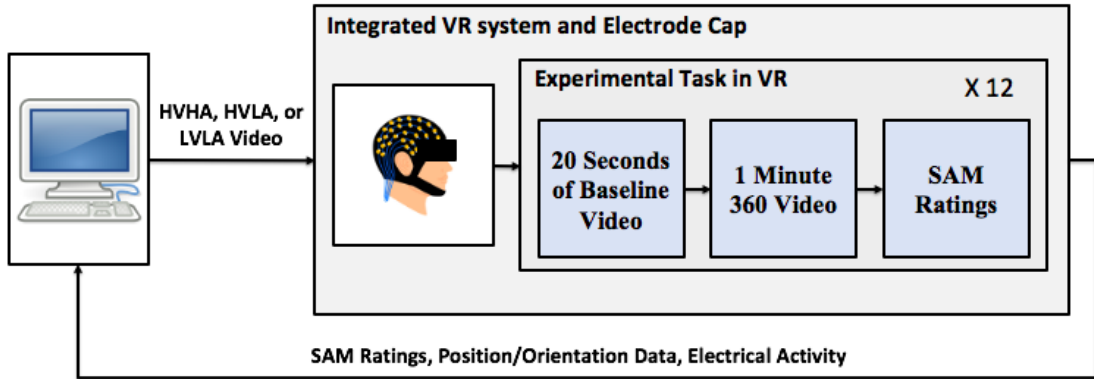
Fig. 12. Diagram of experimental task.

4.3   Data Collection

EEG data were collected using a 32-channel wireless biosignal amplifier (g.Nautilus, Guger Technologies) grounded to location AF3, referenced to the right earlobe, and sampled at 250 Hz. The electrode positions are based on the International 10-20 system as shown in Figure 13. Communication between the VR environment (developed in Unity [60]) and the EEG recording was performed via Lab Streaming Layer (LSL) and recorded using the LSL Lab Recorder application. An EEG amplifier attached to the electrode cap communicates to a base station via Bluetooth. The base station was located approximately 1 meter away from the participant. During the initial data acquisition, the EEG data was bandpass filtered from 0.1-100 Hz and notch filtered from 58-62 Hz.

The position and orientation of the headset, also referred to as the pose, were recorded using the SteamVR plugin at a sample rate of 90 Hz. The position vector is output in the inertial coordinate frame, while the orientation vector consists of Euler angles that describe the rotation between the headset (body) coordinate system and
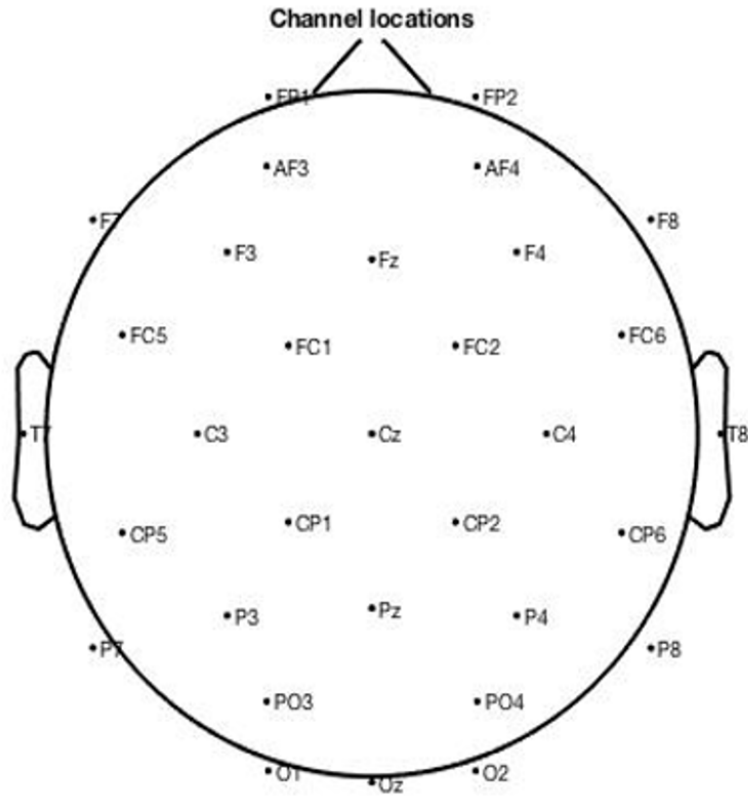
Fig. 13. The EEG cap channel locations using International 10-20 system.

the inertial coordinate system. The HTC VIVE system estimates the pose using fused angle measurements obtained from light house light sweeps and trackers equipped with photodiodes to measure light pulse timings. These angle measurements are also combined with accelerations obtained from an on-board Inertial Measurement Unit (IMU). The accuracy and precision of the HTC position tracking system has been extensively evaluated and further details of the system's pose estimation approach can be found in [61, 62]. In addition to the EEG measurements and pose data, the video ID list sequence, video start and end time stamps, and SAM ratings for each video were recorded for each participant using the LSL.

Video Category	Video Name	Video Description
High Valence / High Arousal (HVHA)	1. Speed Flying 2. Walk the Tight Rope 3. Mega Coaster 4. Through Mowgli's Eyes	1. Viewer follows a speed wing pilot as he glides past mountains 2. Viewer experiences walking a tight rope over a canyon 3. Viewer takes the perspective of riding roller coaster 4. A short film where the viewer observes a conversation between an ape and a bear
Low Valence / Low Arousal (LVLA)	5. Nepal Earthquake 6. Abandoned City 7. Chernobyl 8. Happyland	5. Clip on the effects of an earthquake in Nepal 6. Virtual environment of post-apocalyptic abandoned city 7. Short clip of the effects of Chernobyl nuclear disaster 8. Documentary on a Manila dumpsite
High Valence / Low Arousal (HVLA)	9. Instant Caribbean Vacation 10. Mountain Stillness 11. Pacific Sunset 12. Malaekahana Sunrise	9. Promotional video of Caribbean cruise 10. Atmospheric shots of Canadian Mountains 11. Time-lapse clip of the sunset 12. Viewer observes the sun rising over the horizon at beach
Neutral Video	13. Surrounded by Elephants	13. Viewer has up close experience with elephants in a field

Table 3. Video categories and videos selected from public database with description of each video [7].

CHAPTER 5

AFFECTIVE STATE RATINGS

This chapter examines the resulting subjective SAM ratings across the 30 participants for the 3 categories of the valence-arousal space (HVHA, LVLA, HVLA). The distribution of ratings for each category and the mean intercorrelations over the four SAM scales: valence, arousal, dominance, liking, are presented.

5.1 Analysis of Subjective SAM Ratings

The aggregate SAM ratings for the 3 video categories is explored to determine distribution of the ratings across 30 participants. Each video viewed by participants was grouped by its corresponding video category: high valence high arousal (HVHA), low valence low arousal (LVLA), high valence high arousal (HVHA). The statistics of the SAM ratings were computed to provide a median, minimum, maximum, and standard deviation of each affective state (valence, arousal, dominance, liking). Figures 14 depicts box plots of ratings for the three video categories.

Videos in the HVHA video category had a median of 6.5 for valence, 6 for arousal, 7 for liking, and 6 for dominance. For the HVLA video ratings, valence and liking had a respective medians of 8 and 7, larger than the median of 6 for arousal and 5 for dominance. Videos categorized as LVLA had comparatively low medians ranging for 3-5 for valence, arousal, and liking. The median value for dominance was 7 for LVLA, suggesting the videos are more captivating than the other categories. While the median values correspond to the affective state categorizations as determined by [7], for each categorization it is noted that the range of ratings is consistently large

36

for all affective states.

An analysis of variance (ANOVA) was performed in order to validate videos selected for the experiment and to determine whether valence, arousal, dominance, and liking ratings were systematically different across the three video categories (HVHA, LVLA, HVLA). The MATLAB function *anovan* was used and the inputs included a response vector consisting of ratings from the 30 participants labeled as three groups: video categories, affective states, and participants. Using the results of the ANOVA, a posthoc Tukey-Kramer test was performed using the MATLAB function *multcompare*. Results indicate the population means for valence, liking, and dominance are significantly different across the three video categories. High valence (HV) video categories have a significantly greater mean than low valence (LV) videos. High arousal (HA) video categories also have a significantly greater mean than low arousal (LA) videos. Participants were able to differentiate between high versus low ratings for each category type, however the ranges of the means varied across the four affective states.

The mean intercorrelation of the different scales over the 30 participants (see Table 4) was explored to indicate possible confounds or unwanted effects of fatigue from certain video categories. All intercorrelations were found to be statistically significant, with a large positive correlation between liking and valence (ρ = 0.83). Correlations ranging from ρ = 0.33 to ρ = 0.44 were observed between valence/arousal, dominance/arousal, and between arousal/liking. The correlation between dominance/liking was ranked the second lowest (ρ = 0.21). The lowest intercorrelation was observed between valence and dominance (ρ =- 0.002).
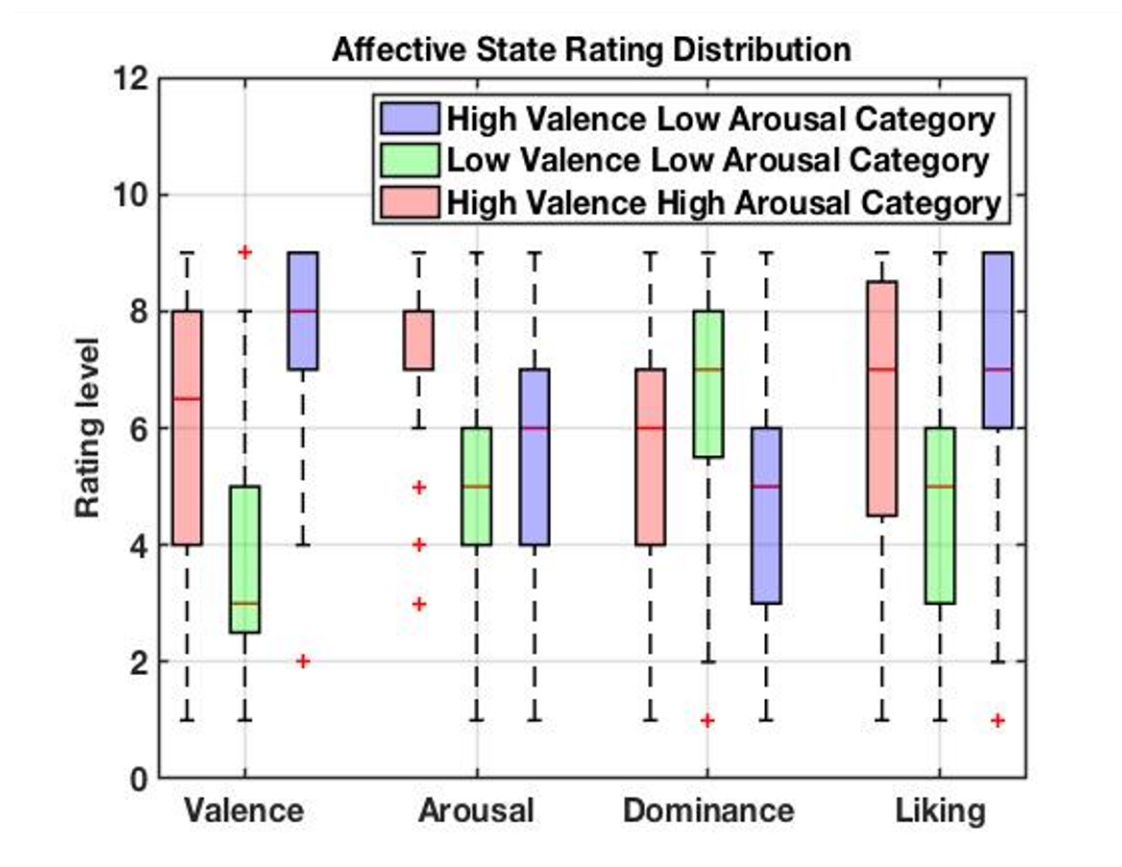
Fig. 14. Boxplot of affective state ratings for HVLA,LVLA,HVHA videos.

5.2   Discussion

The reported affect scores generally aligned with the affective state categorizations HVHA, LVLA, and HVLA. The HVHA category exhibited the highest median rating for arousal. The median ranges from 6-7 for all affective states. This supports the notion that videos which are pleasing are generally rated above 5 for each state. Rating videos aboves a 5 for valence and arousal indicates the participants experienced positive emotion while watching these videos. Videos in the LVLA category had a median of 3 for valence and 5 for arousal. Overall the liking score was lower for LVLA than HVLA and HVHA, indicating videos intended to elicit negative emotion

	Valence	Arousal	Dominance	Liking
Valence	1	0.33*	-0.002	0.83*
Arousal		1	0.44*	0.42*
Dominance			1	0.21*
Liking				1

Table 4. Means of the participant-wise intercorrelations between the scales of valence, arousal, liking, dominance, for all 30 participants. Significant correlations ($p \leq 0.05$) using Fisher's method are indicated by *.

were rated as unpleasant and were generally disliked. Dominance rating signifies the feeling of insignificant versus influential.

For the category of LVLA, the higher dominance median can be correlated to the impactful video excerpts shown to the participant. Videos meant to elicit a calm/relaxed feeling were presented in the HVLA category. Videos presented in this category were generally rated high for valence and liking. The median rating of 8 and 7 for valence and liking suggest these videos elicited a positive/pleasant feeling. However, the median rating for arousal is a 6, indicating that high valence can lead to a higher arousal and liking rating. These videos were grouped to be relaxing and had a median rating of 5 for dominance, which is less than the LVLA and HVHA categories and thus were generally considered to be insignificant, as expected. While the median values correspond to the affective state categorization as determined by [7], it is noted that the range of ratings consistently large for all affective states.

The intercorrelations in Table 4 indicate that there are significant relationships

between most pairs of affect scores except valence and dominance. In terms of valence and arousal, this is likely a function of the prescribed categorizations, with HVHA and LVLA creating inherent correlation. The comparatively large correlation between valence and liking is unsurprising since positive/negative emotions (valence) are generally associated with liking/disliking, respectively. In summary, the affect elicitation was in general successful, though the low valence conditions were partially biased by moderate arousal and liking responses. High scale inter-correlations observed are limited to the scale of valence with those of liking and might be expected in the context of video selection. The rest of the scale intercorrelations are small or medium in strength, indicating the scale concepts were well distinguished by the participants.

CHAPTER 6

EEG ARTIFACT SUPPRESSION

Artifacts can obscure EEG data and limit the analysis of the frequency and time-domain characteristics of the recorded signals; therefore, it is critical to reduce the influence of artifacts in the data. Artifacts can be considered contributions to the EEG signal that are not generated by brain activity. Common sources of physiological artifacts include eye movements, sweating, cardiac activity, and muscle movements [63]. Careful consideration is required when removing components of the EEG signal that are potentially associated with these artifacts, as it could also remove contributions due to brain activity. In this experiment, participants frequently move their heads to view different regions within the VR environment, resulting nonzero headset translational accelerations and Euler angle excursions. This chapter presents a methodology for identifying suppressing components of the EEG signal associated with head movement. This is followed with comparisons of temporal traces and power spectral estimates before and after artifact suppression.

6.1   Data Preprocessing

Data preprocessing is performed prior to applying artifact suppression methods to reduce the influence of noise and other external disturbances. The EEG data was bandpass filtered between 4-47 Hz, which provides an adequate frequency range for analysis of the power bands used for subsequent analysis. A visual inspection of the time-series data is then required to mark noisy channels for removal that contain uniformly large amplitudes or other atypical morphologies. Further analysis of the

power spectral content for the marked channels confirms that these provide unreasonable amplitude contributions relative to the 'clean' channels. Common causes of excessively noisy channels is inadequate contact between the electrode and the participant's scalp due to dried electrolyte gel or an ill-fitting cap, for instance.

After the data is filtered and the channel removal process is complete, the EEG and headset position data are segmented into 12 one-minute trials for each participant. To correct for unrelated stimulus variation in power, the baseline EEG signal from the 20 second inter-trial interval was used.

6.2 Artifact Suppression from Headset Movements

In a VR environment with a wireless EEG headset, participants are able to make naturalistic movements during passive observation. The kinematic characteristics of the movement varies between participant and trial. For example, there are time intervals of small amplitude movements while a participant focuses gaze in a particular region of the virtual environment, or conversely intervals with non-uniform, large amplitude head movement as the participant moves his or her head to change perspectives. Head and body movement artifacts have been observed in previous work to largely influence the frequency content of EEG signals, particularly at frequencies below 20 Hz [64]. Consequently, filtering methods and artifact reduction approaches have been proposed to mitigate these effects [64, 65]. The artifact suppression process described in [65] was selected for application to the EEG data.

Second-order Blind Identification (SOBI) independent component analysis (ICA) is applied to separate the EEG data into 'components' that are maximally statistically independent from one another. SOBI is an extension of ICA method based upon the concept of joint diagonalizaiton of time lagged co-variance matrices and was selected because it is known to be effective at separating neuronal sources and artifacts [66].

The MATLAB function *sobi* was used to apply SOBI ICA to the EEG data of each trial for each participant [67].

The position data was collected with a sampling rate of 90 Hz and, in order to directly compare with the EEG data, the position data was interpolated to 250 Hz. The Spearman correlation coefficient was then computed between each individual component and each axis of the headset (body-axis) acceleration. The headset accelerations were computed numerically from the HTC VIVE inertial position and orientation measurements. The inertial positions were transformed into the headset coordinate frame using a rotation matrix constructed from the roll, pitch, and yaw measurements. The translational accelerations were then computed by taking the second derivative of the headset position in the headset coordinate frame. A threshold was established to specify which components should be retained or removed. Independent components with a Spearman correlation coefficient outside two standard deviations from the mean correlation coefficient were selected for removal. The EEG signal was then reconstructed using only the retained components.

An example of the head movement artifact suppression process is given for participant 11, trial 12. A time trace of the headset translational accelerations and Euler angles are provided in Figure 15 for a time interval of 35-60 seconds. There are large amplitude spikes due to headset acceleration between approximately 43-47 seconds and 55-58 seconds. The amplitude spikes are primarily dominant in the x and z axes. This can be a result of the participant moving his or her head to view different areas within the VR environment. This is supported with the time trace plot of the Euler angles where large amplitude yaw angle excursions begin at 43 and 55 seconds, respectively. The linear increase in yaw angle between approximately 43-47 seconds represent the participant turning his or her head from left to right. The pitch angle may be interpreted as the participant moving between between nose-up (positive)

and nose-down (negative) orientations. The roll angles represent the participant's head tilt either to the right (positive) or left (negative). The roll angle values do not deviate significantly from level. These acceleration time histories were then correlated to the EEG components output from SOBI ICA for the corresponding trial.
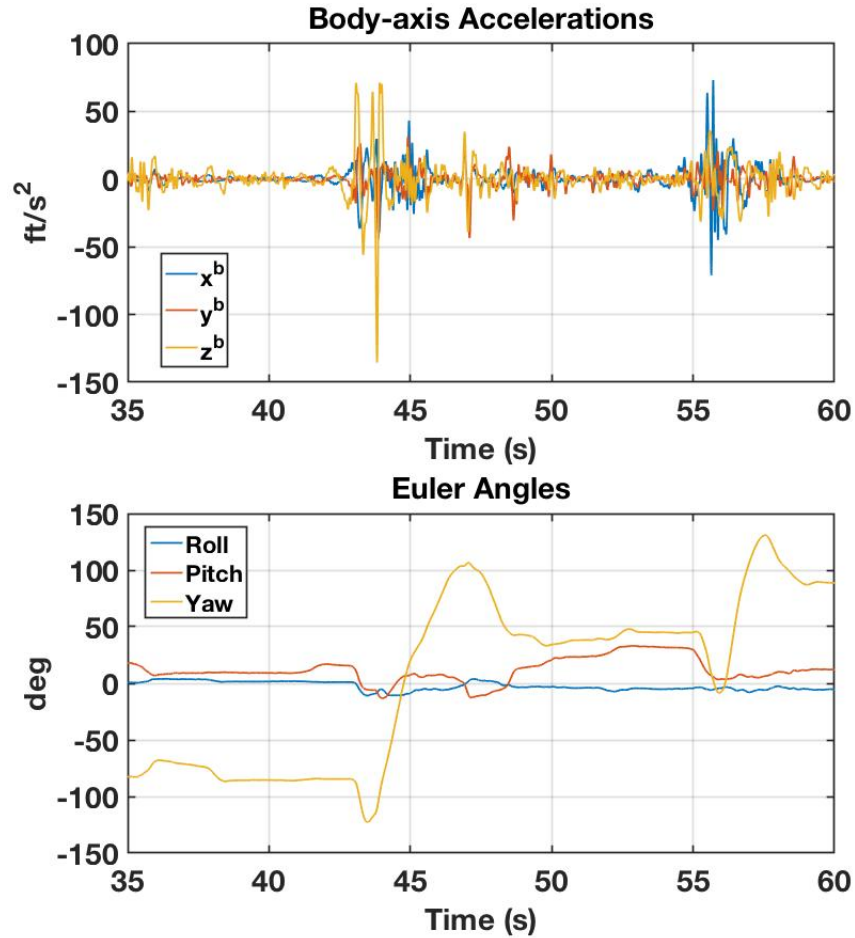
Fig. 15. Headset accelerations for participant 11, trial 12. (top); corresponding Euler angles (bottom).

Figure 16 shows a comparison of the original EEG signal (blue) and reconstructed signal after component removal (orange) for participant 11, trial 12. Each plot shows the time trace of a channel from the frontal (FC6), central parietal (CP5), parietal

(P7), and temporal (T8) regions, respectively. For each channel, several large amplitude spikes between -400 to 250 μV are observed between approximately 43-51 s and 55-58 s. Nominal EEG data is normally bounded between -50 and 50 μV. Visual comparison between the EEG signal amplitudes spikes and the acceleration time history indicate these are potentially an artifact of headset movement. Removal of the components with a Spearman correlation outside two standard deviations from the mean eliminate majority of these large amplitude spikes, as observed in the orange lines. The increased noise observed in the temporal channel (T8) may be due to external disturbance from the headset audio.
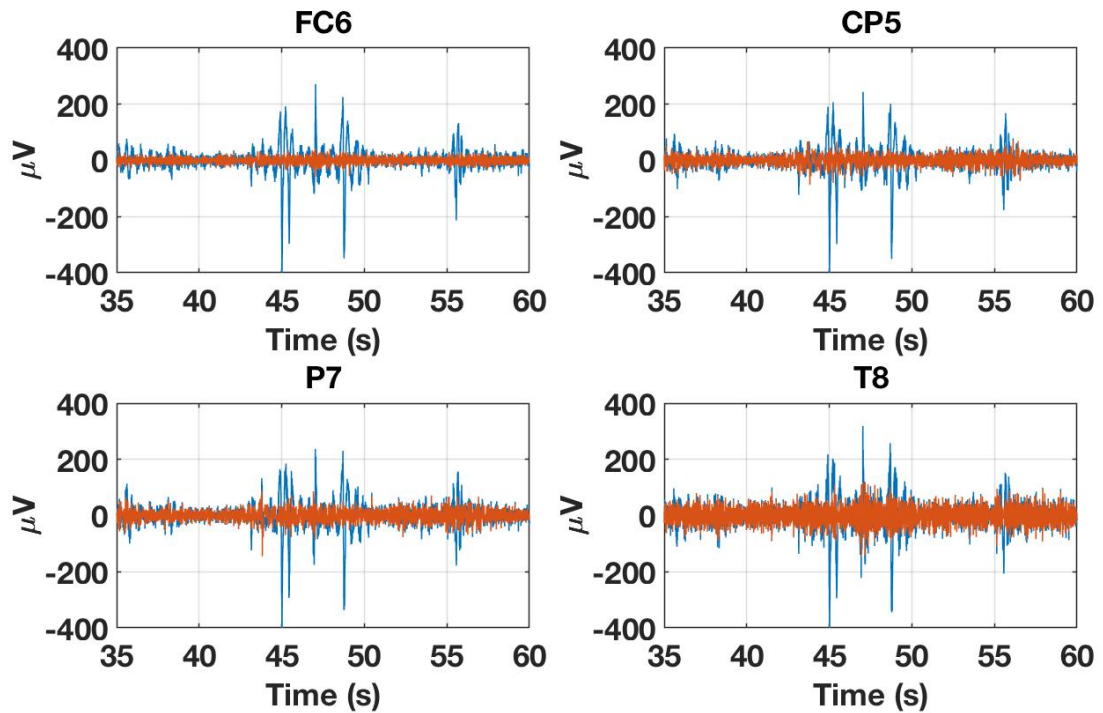
Fig. 16. EEG signal traces from participant 11, trial 12. The blue line represents the original signal and the orange line represents the reconstructed signal after component removal.

Figure 17 depicts the power spectra of the four channels before and after head

movement artifact suppression. The blue line in each plot represents the original signal without artifact suppression, and the orange line represents the power spectra after completion of SOBI ICA. The power spectra estimates were computed using the *pwelch* MATLAB function with a window size of 5 seconds and 0 overlap. The most significant decrease in power before and after artifact suppression are observed in the theta (4-7 Hz) and alpha (8-14 Hz) band. Less effects of the artifact suppression method are seen in higher frequency bands, as it is primarily expected to influence power of frequencies below 20 Hz due to gross movements [64].
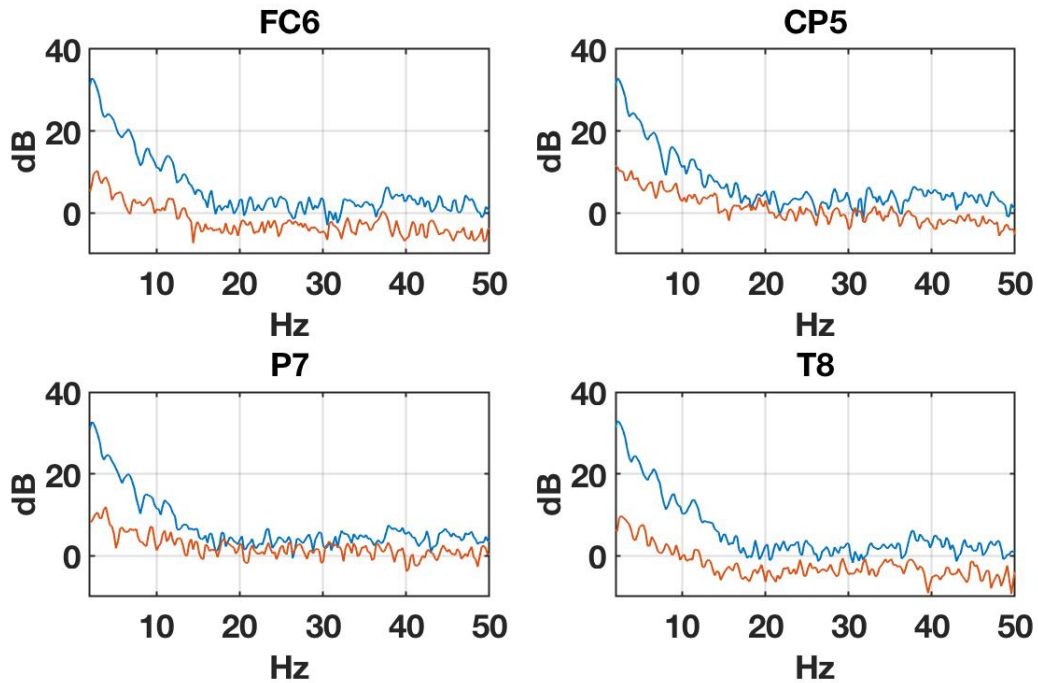
Fig. 17. Power spectra of 4 representative EEG channels before artifact suppression (blue) and after artifact suppression (orange).

In summary, it is critical to address large amplitude spike artifacts observed in the EEG signal using artifact suppression techniques prior to subsequent analysis since the different video categories may inherently elicit different degrees of head

movements. The proposed method was found to effectively reduce the influence of head movement artifacts in the EEG signal. The magnitude of other artifacts such as eye blink artifacts may also be reduced after application of the method. This approach is applied to the data prior to the analysis presented in subsequent chapters.

CHAPTER 7

CORRELATIONS BETWEEN EEG SPECTRAL FEATURES AND SAM RATINGS

7.1 Bandpower Correlation with Affective State

The power spectral density of the EEG signal for each trial and baseline were computed using Welch's method with window size of 10 seconds and an overlap of 50%. The baseline power was subtracted from the trial power, yielding the change of power relative to pre-stimulus period. The net changes of power spectral density (PSD) were then averaged over four frequency bands: θ (4-7 Hz), α (8-13 Hz), β (14-29 Hz), and γ (30-47 Hz).

Spearman's correlation coefficient, ρ, between the power changes and the subjective ratings was computed along with the p-values for positive and negative correlation tests. This was done for each participant separately and, assuming independence, the 32-resulting p-values per correlation direction (positive/negative), frequency band, and electrode were combined to one p-value using Fisher's method [68, 69]. Figure 18 depicts the results of the correlation analysis between participant ratings and EEG frequency power. The analysis suggests that brain activity from different regions of the scalp can be related to the subjective affective state evaluation of the participant along the axes of arousal and valence, and to their preference for the videos. The large number of tests computed may lead to an increase in false positives therefore a Bonferonni correction criteria (0.05/32) is applied and the resulting highly significant (p≤0.0016) correlations are discussed. A comprehensive list of the effects can be found in Tables 5-7.

48
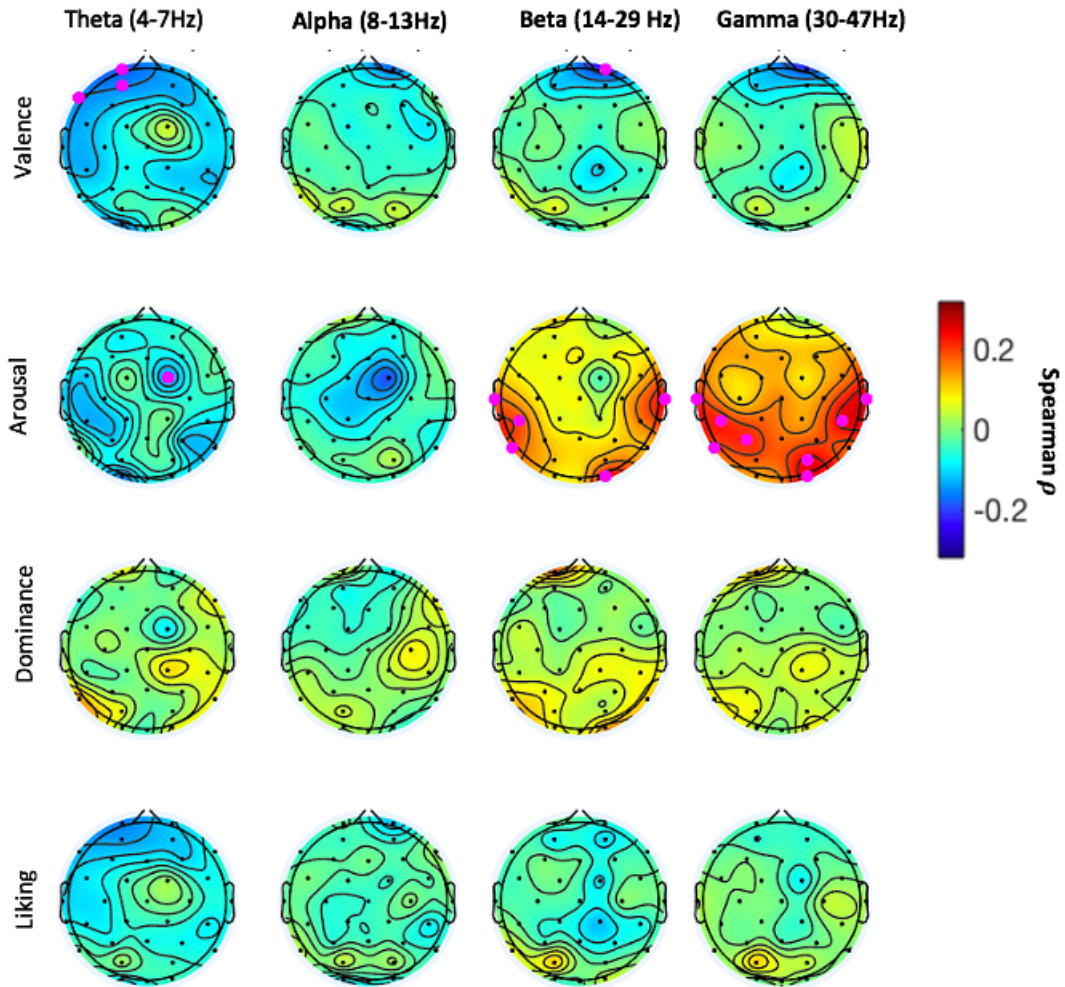
Fig. 18. The mean Spearman correlations over 30 participants between valence, arousal, dominance, and liking with the power in the frequency bands of θ (4-7 Hz), α (8-13 Hz), β (14-29 Hz), and γ (30-47 Hz), respectively. The pink highlighted electrodes correlate significantly ($p \leq 0.0016$) with ratings.

Valence exhibited the strongest inverse correlation, and correlates were found in theta and beta frequency bands. Theta valence showed correlates in the left frontal region, and has a strong negative correlation. Beta has a correlate over the frontal parietal region. For arousal, negative correlations existed in the theta and alpha bands. A central alpha power decrease was observed in the alpha band. Positive arousal correlations are observed in the higher frequency bands. The correlates present in the

beta band are located at the temporal, occipital,and parietal regions. The gamma arousal region contains the most correlates.

For the dominance affective state, no significant correlates are observed. The higher frequency bands depict a power increase towards the frontal region and parietal and temporal regions.

Liking depicted a similar trend to valence, but no significant correlate was observed in theta liking when a Bonferonni correction is applied. Theta valence and theta liking have negative correlation towards the frontal region. Overall, both alpha valence and alpha liking follow similar trends with an slight increase in power towards the occipital region.

Valence and arousal affective states have correlates present with a p-value $\leq.0016$. Valence has significant electrodes present in the frontal region for theta and beta. The arousal rating contains significant electrodes in the temporal region in both beta and gamma bands; however, most of the significant electrodes are present in the central parietal region for the gamma band. Similar to valence, liking depicts the same correlation trend as theta valence. The distribution of p-values in Tables 5-7 also suggests a large distribution of values between the max and minimum value for each significant electrode.

Theta				
	Electrode	$\rho-$	$\rho+$	$\bar{\rho}$
Valence	FP1	-0.76	0.44	-0.18
	AF3	-0.76	0.62	-0.16
	F7	-0.87	0.43	-0.15
Arousal	FC2	-0.67	0.31	-0.14

Table 5. Electrodes exhibiting significant correlation in theta band (4-7Hz) and scale (*p $\leq$ 0.0016.)

Beta				
	Electrode	$\rho-$	$\rho+$	$\bar{\rho}$
Valence	FP2	-0.48	0.47	-0.18
Arousal	T7	-0.34	0.85	0.17
	P7	-0.43	0.79	0.18
	O2	-0.28	0.71	0.22
	T8	-0.48	0.58	0.24

Table 6. Electrodes exhibiting significant correlation in beta band (14-29Hz) and scale (*p $\leq$ 0.0016.)

Gamma				
	Electrode	$\rho-$	$\rho+$	$\bar{\rho}$
Arousal	T7	-0.66	0.69	0.16
	CP5	-0.50	0.74	0.18
	P3	-0.53	0.57	0.12
	P7	-0.44	0.77	0.18
	PO4	-0.36	0.77	0.14
	O2	-0.30	0.70	0.22
	T8	-0.38	0.63	0.23
	CP6	-0.50	0.70	0.16

Table 7. Electrodes exhibiting significant correlation in gamma band (30-47Hz) and scale (*p $\leq$ 0.0016.)

7.2 Discussion

The Spearman correlation provides insight to which frequencies and regions of the brain correlate with affective states. Correlations observed partially concur with observations made in other studies exploring neurological correlates of affective states. Frontal parietal and anterior frontal electrodes are significant in valence in the theta and beta bands. These results can be due to frontal lobe activation during evoked pleasurable emotion [2]. The location of correlates over the occipital region in beta and gamma arousal, suggest a relative deactivation, or top-down inhibition due to the

participants focusing on pleasurable sound [70]. Increased beta power over left and right temporal sites in arousal is also associated with positive emotional self-induction and external stimulation [71]. Koelestra et. al also did not depict any correlates in dominance [68], this can be due to the rating category being too subjective and vague for a group of participants.

In general, the distribution of valence and liking correlations shown in Figure 18 are similar to previous affective studies [68]. The similarity between the two groups may be a result of the high intercorrelations of the self-assessment ratings. A video that induces pleasant feeling will generally be liked, and rated similarly. Positive correlation of arousal, particularly in the gamma band, emanating from the anterior temporal region have also been reported [68]. However, electromyographic (EMG) activity is also known to be prominent in the higher frequencies over anterior and temporal electrodes, particularly as a result of emotive facial expressions [72]. While EMG was not obvious during visual inspection of the present signals, modulations due to more subtle muscle tension cannot be ruled out. Due to large distribution of ratings and the presence of p-values hovering around 0, high inter-participant variability is expected. Therefore for modeling purposes a participant-specific modeling approach is taken, rather than single model for all participants.

CHAPTER 8

AFFECTIVE STATE ESTIMATION USING EEG

This chapter presents modeling approaches and results for affective state estimation. The primary objective of the modeling efforts is to demonstrate the feasibility of predicting a wide range of affective state levels using EEG features. Spectral features were computed from EEG signals for each electrode and utilized as explanatory variables to estimate the subjective SAM ratings. The model performance of a stepwise regression model and a binary Support Vector Machine (SVM) are presented and discussed. The application of these models is to support the development of future immersive VR experiences by providing human affective state feedback in a closed-loop control structure.

8.1   Selection of Data for Training and Testing

As with the majority of BCI research, the presence of high inter-participant variability justifies a participant-specific modeling and classification approach. Thus, affective state models were developed separately for each individual participant; each participant completed 12 one-minute experimental tasks where a rating was provided for each affective state. The EEG signals for each trial were segmented into 12 five-second intervals, this results in 144 segments per participant. The power spectral density was estimated for each of these five-second segments individually using the MATLAB *pwelch* function.

A four-fold cross validation technique was applied to evaluate various combinations of data utilized for model training and testing. A new model is developed

and evaluated at each fold. For example, in the first fold, the power bands for nine randomly selected segments from each trial were input for model training, while the power bands of the three remaining segments were withheld to assess the model classification accuracy. For a given fold, this results in 108 segments (12 trials x 9 segments) utilized for model training, and 36 (12 trials x 3 segments) segments for model testing. In the subsequent folds the process repeats; however, a different set of three segments per trial are withheld from model testing. By the fourth fold, each segment has been utilized for either model testing or training. Figure 19 provides a conceptual representation of the four fold cross validation process.
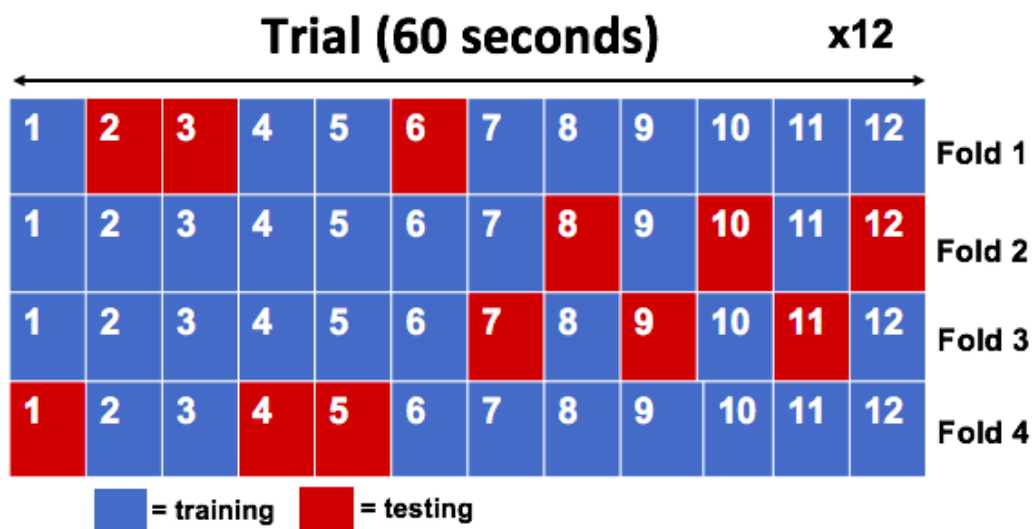
Fig. 19. Conceptual representation of k-fold cross validation process.

8.2 Model Development using Stepwise Regression

Stepwise regression was applied to predict each affective state rating as a function of the measured brain electrical activity during each trial. A major challenge associated with modeling human affective states is that the model structure and significant explanatory variables have not been previously established. Additionally, there are

a large number of potential explanatory variables given that there are 32 electrodes, resulting in an over-determined problem. Therefore, a stepwise regression routine was utilized to identify statistically significant model terms.

The pool of candidate explanatory variables consisted of the power bands for each electrode, resulting in 128 potential terms (32 channels x 4 power bands). The candidate terms were restricted to the first order, neglecting cross and higher order combinations of terms. Future work may explore the presence of nonlinear terms on model performance. The response variables consisted of the affective state ratings, with each rating ranging from one to nine in intervals of one. The MATLAB function *stepwisefit* was used to automate the stepwise regression process. The process begins with a forward selection procedure where no terms are initially in the model structure. Successive steps are then taken to include or discard terms into the model based on several statistical metrics. At each step, the p-value and coefficient of determination (R^2) are evaluated to determine whether a term should remain in the model. The criteria requires that each term retained in the model has a p-value less than 0.05 and increases the R^2 by a minimum of 0.5%.

To assess the impact of artifact suppression on model performance, models were developed using EEG spectral data before and after artifact suppression was applied. To further examine the potential contribution of movement artifacts, an additional model was developed using exclusively headset acceleration data.

There are several parameters than can be modified when estimating the power spectral density of the signal, such as window size and overlap. To further optimize the model with respect to the dynamics of affective state, the influence of these parameters on model accuracy was explored using the R^2 as a metric to assess performance. The average R^2 across participants was computed using a 1 and 5 second window size and a 0% and 50% overlap. These results for the model development and model prediction

stages are provided in Table 8. It can be seen that increasing the window size from 1 second to 5 seconds with 0% overlap increases the model prediction R^2 by 10% for valence. Similar increases are observed for arousal, dominance, and liking. A 1 second window size resulted in the lowest average R^2 values for model development and model prediction. A 5 second window with 50% overlap did not have a significant impact on the average R^2 when compared to the results using a 1 second window size.

Window Size (Sec) Overlap (%)	Average R^2 Model Development			Average R^2 Model Prediction		
	1 sec 0 %	5 sec 0%	5 sec 50%	1 sec 0%	5 sec 0%	5 sec 50%
Valence	73%	82%	76%	56%	66%	57%
Arousal	76%	81%	77%	59%	66%	65%
Dominance	77%	81%	77%	62%	67%	63%
Liking	78%	81%	79%	61%	69%	62%

Table 8. Influence of Welch's Method parameters on average R^2 computed by stepwise regression.

The distribution of the R^2 values for the models developed using a 5 second window with 0% overlap is given in Figure 20. The median R^2 value is approximately 70% for each rating category after artifact suppression. Artifact suppression resulted in 10-15% increase in median R^2, indicating that the movement artifacts were likely masking relevant EEG features. Furthermore, the accelerometer data alone was not capable of estimating the SAM ratings, providing further evidence that EEG features

are primarily contributing the predictive power. These results show that the models can predict the general trend of the affective state ratings, the models perform similarly across participants for each affective state, and the artifact suppression provides some performance benefit.
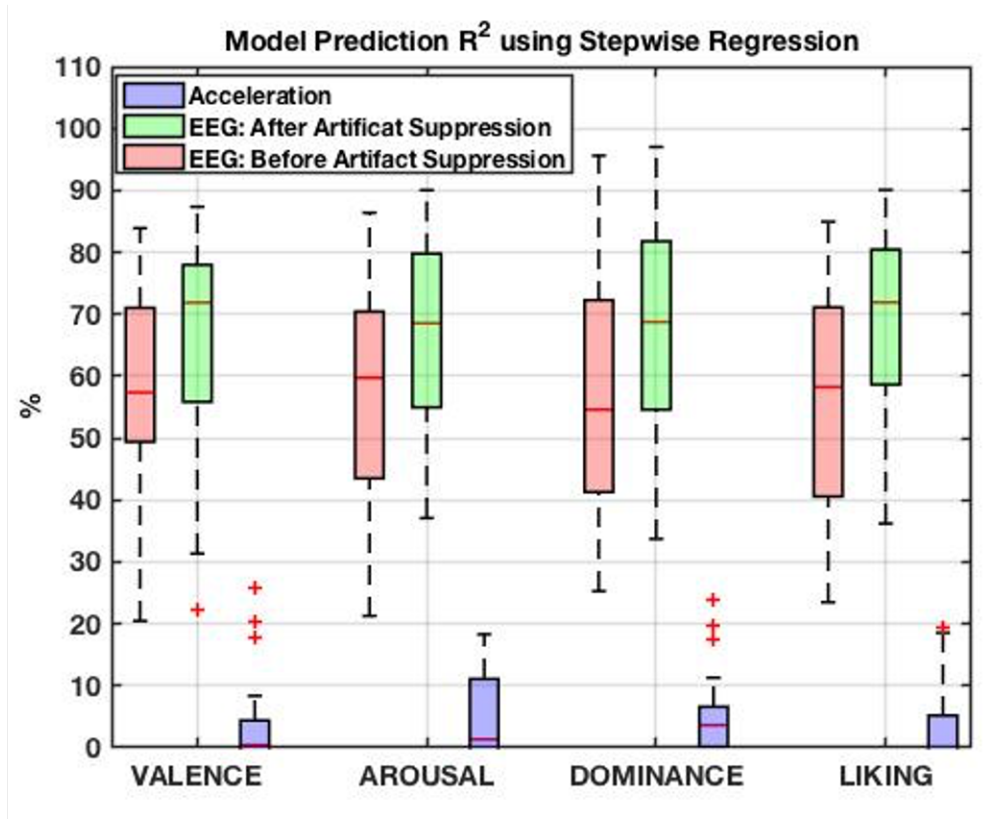

Fig. 20. Comparison of average R^2 for affective state models developed using stepwise regression. The power bands were computed using the EEG signals with 5 second window sizes and 0% overlap.

In order to better assess the contributions of the electrodes to the models and the consistency across participants, an analysis of the most commonly selected electrodes across participants was performed. Electrodes that were included in the models for at least three of the four folds for a given participant were denoted as relevant for the participant and recorded. The resulting relevant electrodes were aggregated across

the 30 participants as depicted in Figure 21.

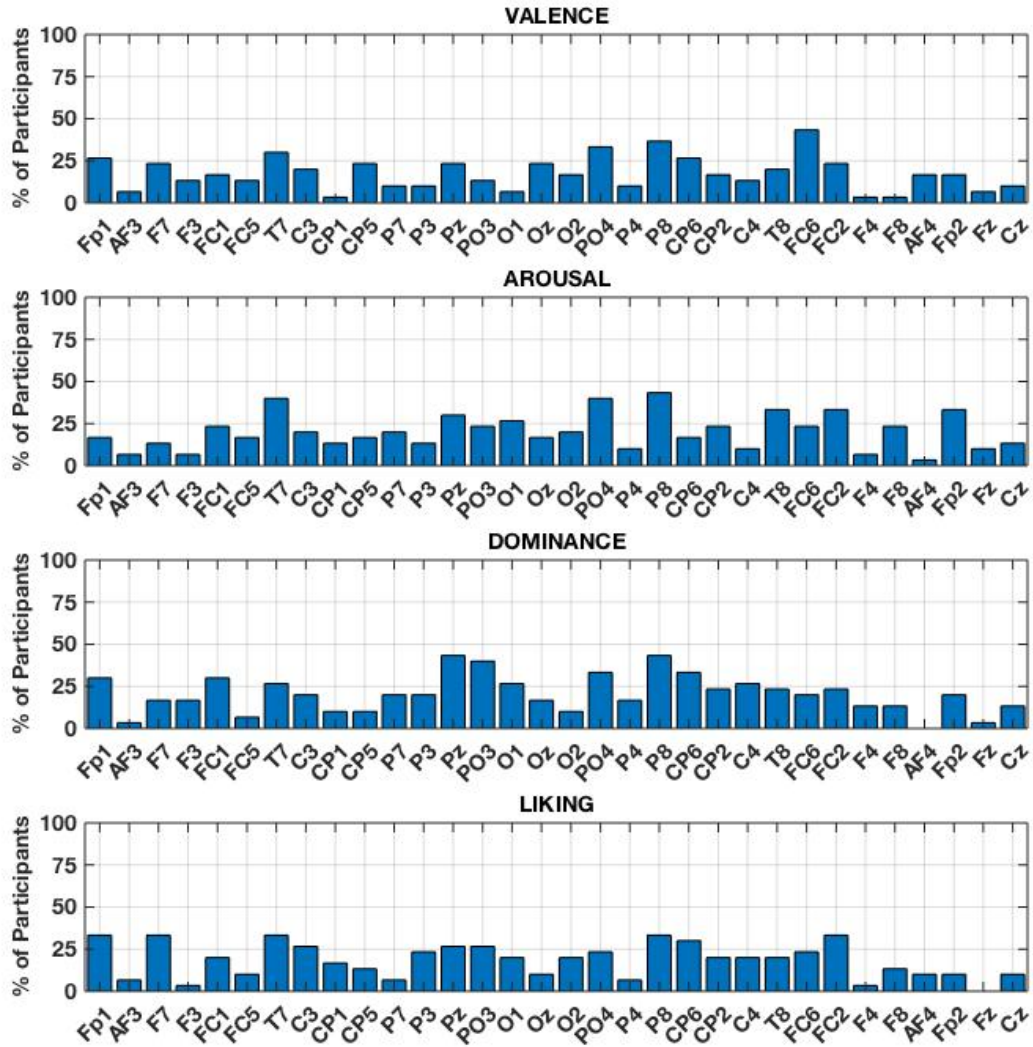Fig. 21. Percentage of participants for which each electrode was determined to be relevant to the stepwise regression modeling for each affective state.

Electrodes P04, P8, CP6, and T7 were included in greater than 20% of the affective state models, indicating that the combination of these areas is useful for affective state estimation. All these electrodes were found to be statistically significant

electrodes in the arousal gamma band. Electrode T7 exhibited significant correlation between rating and all three frequency bands: theta, beta, gamma (as discussed in Chapter 7). CP6 and T7 were found to be statistically significant in other data sets studying affective state estimation [1]. PO4, P8, and CP6 are located over the parietal lobe, which processes information associated with sensory integration. This region of the brain may be activated due to the visual stimuli often portraying first-person movements through the scenes. T7 is located on the temporal lobe, which processes sensation of sound, sight, and touch. However, this location is also prone to muscle tension due to facial expressions or jaw clenching. Further analysis into the relevant electrodes for individual affective states may provide guidance in future modeling efforts to which electrodes are significant for a given affective state, particularly in experiments with limited electrode availability.

8.3 Affective State Estimation using Binary Support Vector Machine

For recent research on EEG for estimating affective state, practical performance metrics such as the minimum level of fidelity or accuracy needed have not been defined. For example, while the capability to predict a high resolution of rating levels is desirable, it might not be necessary for a particular application. Alternately, it may be more useful to simply distinguish between low and high levels of an affective state. Therefore, the SAM ratings were thresholded to low and high and the feasibility and performance of a binary Support Vector Machine (SVM) classifier was evaluated.

The SAM ratings (1-9) were transformed into binary ratings (0 or 1) using a threshold determined by the mean rating for each individual participant and affective state, respectively. The four-fold cross validation approach described in the previous sections was applied to define the training and testing data sets. The dimension of the explanatory variable matrix was reduced using the features identified from

the previous stepwise regression process. These features were input into the SVM algorithm for model development. The explanatory variable matrix and response variable vector were input to the MATLAB $fitcsvm$ function for model training, which uses a Gaussian kernel. Similar to the regression models, SVM models were created using EEG spectral bands before and after artifact suppression. An additional classifier was developed using exclusively headset acceleration data for comparison to the EEG-based classifiers.

Because the number of observations of each binary class (i.e., number of high and low ratings) was generally imbalanced and varied between participants, the sensitivity and specificity were computed in addition to model classification accuracy. These performance metrics are described using the terms based on the confusion matrix shown in Table 9.

		Truth		
		High Level Rating	Low Level Rating	Total Number
Test Result	Positive (number)	A True Positive	B False Positive	$T_{Test\ Positive}$
	Negative (number)	C False Negative	D True Negative	$T_{Test\ Negative}$
		T_{High}	T_{Low}	Total

Table 9. Conceptual representation of a confusion matrix.

Using the terms in Table 9, the model classification accuracy is defined in Equation 8.1 as a function of the number of true positives, A, false positives, B, false

negatives, C, and true negatives, D.

$$accuracy = \left(\frac{A + D}{A + B + C + D} \right) \tag{8.1}$$

The sensitivity metric describes the probability that a rating will be labeled a high rating out of the ratings that are defined as high. The calculation for sensitivity is given in Equation 8.2 as a function of the number of true positives, A, and the false negatives, C.

$$sensitivity = \left(\frac{A}{A + C} \right) \tag{8.2}$$

The specificity metric describes the probability that a rating will be labeled a low rating out of the ratings that are defined as low. The calculation for specificity is given in Equation 8.3 as a function of the number of false positives, B, and true negatives, D.

$$specificity = \left(\frac{D}{D + B} \right) \tag{8.3}$$

The distribution of the average model classification accuracy across participants for each affective state are depicted in a box plot in Figure 22. The median value after artifact suppression ranges between approximately 82-90% across affective states. Valence and dominance have the greatest median classification accuracy, followed by liking and arousal. Dominance has the smallest range which is between approximately 70-100%, while arousal has the largest range between approximately 50-98%.

Similar to the stepwise regression results, this indicates that model performance is generally consistent across affective states; however, the effects of artifact suppression are less prominent. In particular, using the headset acceleration data the model performs at approximately 60% for each affective state. Because the linear stepwise regression model was not able to fit the accelerometer data to the SAM scores while nonlinear SVM achieved above-chance performance with the accelerometer data, this

indicates that there is a nonlinear mapping of the accelerometer data that can provide predictive power. Regardless, the superior performance of the models using EEG features suggests that these features provide additional predictive power.
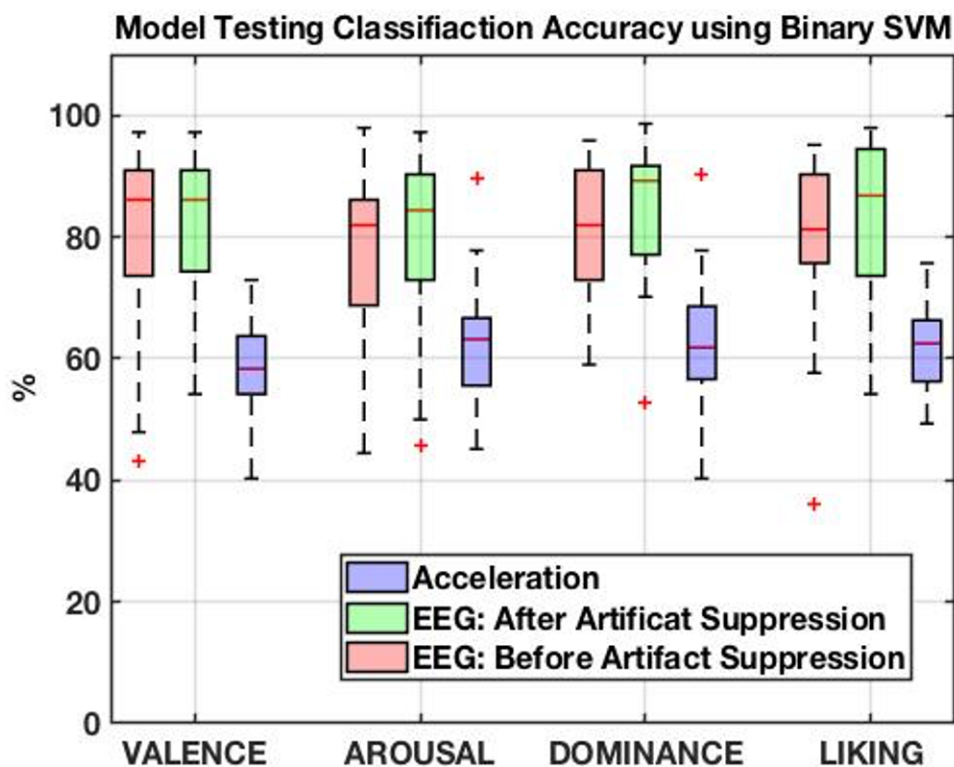

Fig. 22. Comparison of classification accuracy for affective state models developed using support vector machine (SVM). The power bands were computed using the EEG signals with 5 second window sizes and 0% overlap.

To verify whether the classification accuracy results are not simply due to having imbalanced classes, specificity and sensitivity are examined. The distribution of the average sensitivity (top) and specificity (bottom) across each participant and affective state are given in Figure 23. The median percentage for both metrics ranges between 82-86%. This confirms that the models will predict a high rating correctly with the similar probability that it would correctly predict a low rating, for all affective

62

states. Arousal has the largest sensitivity range between approximately 42-100%, while dominance has the smallest sensitivity range between 60-100%. Liking has the highest specificity range from approximately 55-100%, and dominance has the small range between 70-100%. There are significant outliers in valence and arousal average specificity probabilities, reaching as low as 33%.

In summary, the modeling results using a binary SVM classification approach showed reasonable prediction capabilities, which may have advantages over regression models for certain applications. Analysis of the classification accuracy, sensitivity and specificity probabilities show that the model can predict low and high ratings with similar performance. However, the SVM revealed that the accelerometer data can provide predictive power that was not evident from the linear regression model. Future testing with a closed-loop BCI application should be performed to assess if the resulting levels of accuracy are adequate for practical use.
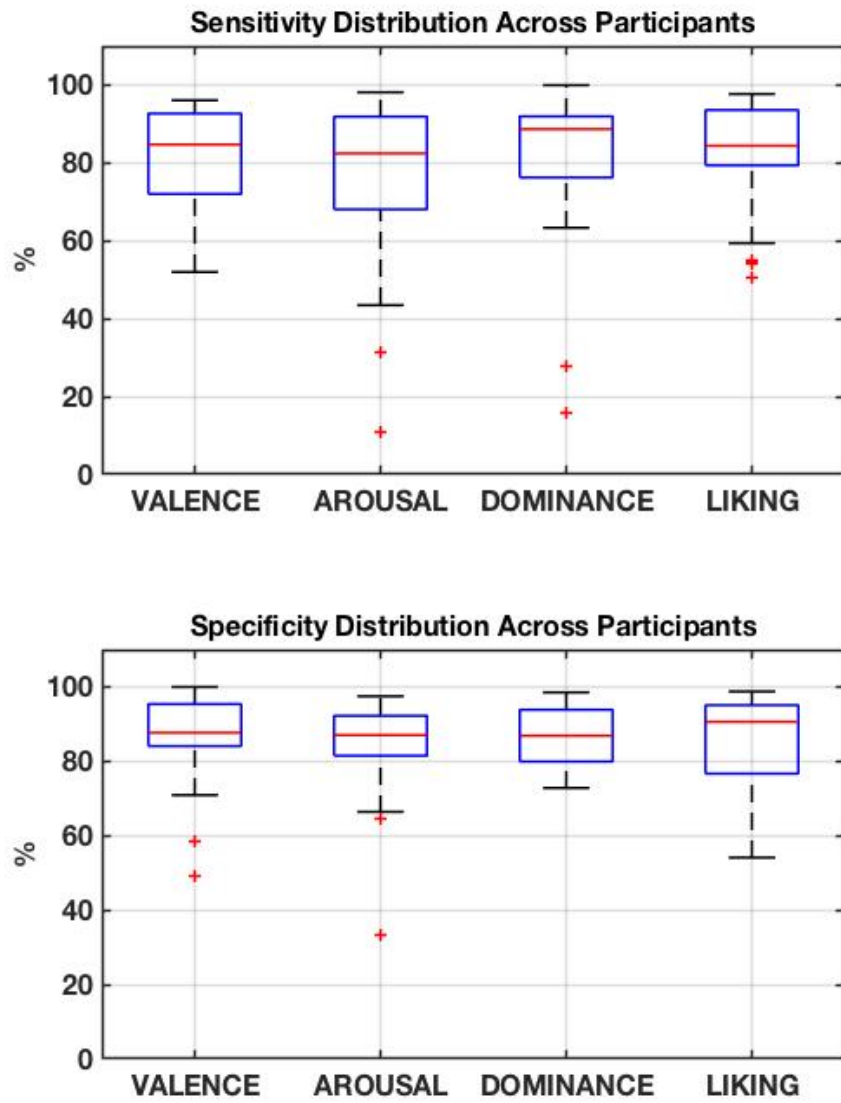
Fig. 23. Average sensitivity and specificity probabilities across the participants for each affective state.

CHAPTER 9

CONCLUSION

The main objective of this thesis was to characterize and estimate affective states induced by 360° virtual environments using scalp-recorded electrophysiological activity. This chapter concludes the thesis with main contributions, limitations, and possible future directions for this research.

9.1 Main Contributions

The results of this thesis demonstrate that subjective self-assessment ratings of a participant's affective state can reliably be estimated via EEG. A novel experimental paradigm and setup was created that involved the integration of a wireless electrode cap, virtual reality headset, and the development of VR presentation and control software.

Because the participant's task generated unavoidable motion artifacts in the EEG, a novel treatment of SOBI ICA was performed to suppress these artifacts. This was novel in the sense that prior studies were based on freely moving EEG, whereas the present study is largely confined to rotational movements, requiring different representations of the accelerometer data. The method was deemed effective based on visual observation of the cleaned EEG an the resulting power spectra.

Using the traditional theta, alpha, beta, and gamma bands of EEG, two novel models were developed to estimate the affective state ratings. Using statistically significant EEG features, a stepwise regression routine was applied yielding models for estimating a rating of 1-9 for each affective state. The resulting R^2 values across

each affective state ranged from 66%-69%. By simplifying the problem to estimate thresholded high versus low affective state ratings, a binary SVM classifier produced a classification accuracy across all affective states that reached above 80%, with performance supported by sensitivity and specificity analyses. Lastly, the highly relevant electrodes based on the affective ratings and the affective state estimation models were aggregated across participants to provide insights to the key contributing brain areas.

9.2 Limitations and Future Work

There are several possible future directions of this research. While artifact suppression methods presented in Chapter 6 appear to effectively suppress head movement artifacts in EEG, no method exists that can definitively separate EEG and electrocardiogram (EMG) or movement artifacts occupying the same spectral frequency bands. Future experimental design should include recording facial muscle EMG activity along with EEG to provide insight on artifacts caused by facial expressions. Additionally, it would have been beneficial to include additional accelerometers to better differentiate head rotations from chair rotations. Ultimately, the combination of EEG, EMG, and movement information would likely improve modeling results, which may more realistic and practical for end-user applications. Due to the subjective nature of the task and varying style and content of the available 360 videos, further analysis needs to be performed on the impact of the specific video stimuli (e.g., motion, audio, realism, optical flow, etc.) on EEG.

This research explored two modeling techniques that are commonly used in EEG affective state research. Other EEG features and modeling approaches should be explored for the possibility of increasing affective state estimation performance. Since the content of the videos and participant's attention changed over time, and that

the objective is to ultimately develop a closed-loop BCI based on affective state, it would be useful to assess the participant's affective state more frequently during the videos rather than once at the end of each video. A limitation to using a 1-9 scale for self-assessment presents a bias due to the scale being participant-dependent, another form of assessment or normalization across participants could also be explored. For the purpose of this experiment, the fourth quadrant of the Russel Complex was not evaluated (Low Valence/High Arousal) and including this quadrant could provide more insight regarding the affective states. Ultimately, the affective state estimation modeling techniques explored could be implemented to provide closed-loop biofeedback to allow user's to alter the virtual reality interactions according to their affective state. This closed-loop scenario introduces various challenges as appropriate feedback and altered affective state dynamics must be investigated in detail.

This study provides a framework for using virtual reality to evoke and estimate affective states in a controlled laboratory environment, and the results obtained provide important insights for future EEG-based affective state estimation studies in VR.

Appendix A

ABBREVIATIONS

BCI   Brain-Computer Interface

EEG   Electroencephalogram

EMG   Electromyogram

HVHA High Valence High Arousal

HVLA   High Valence Low Arousal

BSS   Blind Source Separation

LDA   Linear Discriminant Analysis

LSL   Lab Streaming Layer

LVLA   Low Valence Low Arousal

MSQ   Motion Sickness Susceptibility Questionnaire

PSD   Power Spectral Density

SAM   Self Assessment Manikin

SOBI ICA Second-Order Blind Independent Component Analysis

SVM   Support Vector Machine

VR   Virtual Reality

Appendix B

MOTION SICKNESS SUSCEPTIBILITY QUESTIONNAIRE (MSSQ)

Your childhood experience only (before 12 years of age), for each of the following types of transport or entertainment please indicate

1. **As a child (before age 12),** how often you **felt sick or nauseated** (tick boxes)

	Not Applicable - Never Traveled	Never Felt Sick	Rarely Felt Sick	Sometimes Felt Sick	Frequently Felt Sick
Cars					
Buses or Coaches					
Trains					
Aircraft					
Small Boats					
Ships, e.g. Channel Ferries					
Swings in playgrounds					
Roundabouts in playgrounds					
Big Dippers, Funfair Rides					
	t	0	1	2	3

Your experience over the last 10 years (approximately), for each of the following types of transport or entertainment please indicate

2. **Over the last 10 years,** how often you **felt sick or nauseated** (tick boxes)

	Not Applicable - Never Traveled	Never Felt Sick	Rarely Felt Sick	Sometimes Felt Sick	Frequently Felt Sick
Cars					
Buses or Coaches					
Trains					
Aircraft					
Small Boats					
Ships, e.g. Channel Ferries					
Swings in playgrounds					
Roundabouts in playgrounds					
Big Dippers, Funfair Rides					
	t	0	1	2	3

Appendix C

SELF ASSESSMENT MANIKIN (SAM) SCRIPT

The Self-Assessment Manikin(SAM) scale consists of ratings 1-9. A rating of 1 is represented by the left most block for each category, while nine is represented by the right most block.

Valence - Negative(1) vs. Positive(9)

Valence is the value associated with a stimulus as expressed on continuum from unpleasant to pleasant. The first picture shows a person who is distressed-relevant experiences include irritation, disgust, defeat or crisis. The last picture shows an individual who is elated- relevant experiences could include fun, delight, happiness, relaxation, satisfaction. The remaining pictures depict intermediate states.

Arousal - Low(1) vs. High(9)

Arousal is the state of excitement or energy linked to an emotion. The first picture shows an individual who is very calm, almost sleeping-relevant states could include relaxation, tranquility, meditation, boredom or laziness. The last picture shows an individual who is bursting with arousal-relevant states could include excitation, euphoria, excitement, rage, or agitation.

Dominance - Low (1) vs. High(9)

Dominance refers to having influence or control. The first picture shows an individual who feels a lack of control and agency- relevant states could include subordination, intimidation, subjugation, withdrawal, or resignation. The last picture depicts a person who is dominant and in control of the situation- relevant states include control, influence, being important, dominant, recognized, or decisive.

Liking - Dislike(1) vs. Like(9)

Liking refers to a measure of degree of attraction or repulsion.The first thumbs down represent dislike. The last thumbs up represents dislike. The remaining pictures depict intermediate states.

LINKS TO VIDEOS USED IN EXPERIMENT

Instant Caribbean Vacation

`https://www.youtube.com/watch?v=CDfsFuDuHds`

Mountain Stillness

`https://www.youtube.com/watch?v=aePXpV8Z10Y`

Pacific Sunset

`https://www.youtube.com/watch?v=ka3qzz-oa-A`

Malaekahana sunrise

`https://www.youtube.com/watch?v=-bIrUYM-GjU`

Speed Flying

`https://www.youtube.com/watch?v=g6w6xkQeSHg`

Walk the tightrope

`https://www.youtube.com/watch?v=JtAzMFcUQ90`

Mega coaster

`https://www.youtube.com/watch?v=-xNN-bJQ4vI`

Through Mowglli eyes

`https://www.youtube.com/watch?v=bUiP-iGN6oI`

Nepal earthquake

`https://www.youtube.com/watch?v=5tasUGQ1898`

Abandoned city

`https://www.youtube.com/watch?v=LbxAc784608&spfreload=5`

Chernobyl

https://www.youtube.com/watch?v=VbvgQvH3j34

Happyland

https://www.youtube.com/watch?v=lWDqeWFJdp4

Surrounded by elephants

https://www.youtube.com/watch?v=mlOiXMvMaZo

REFERENCES

[1] S. Koelstra et al. "DEAP: A Database for Emotion Analysis ;Using Physiological Signals". In: *IEEE Transactions on Affective Computing* 3.1 (2012), pp. 18–31. DOI: `10.1109/t-affc.2011.15`.

[2] Paul Ekman and Richard J. Davidson. "Voluntary Smiling Changes Regional Brain Activity". In: *Psychological Science* 4.5 (1993), pp. 342–345. DOI: `10.1111/j.1467-9280.1993.tb00576.x`.

[3] Benjamin Blankertz et al. "The Berlin Brain–Computer Interface: Non-Medical Uses of BCI Technology". In: *Frontiers in Neuroscience* 4 (2010). DOI: `10.3389/fnins.2010.00198`.

[4] J.R. Wolpaw et al. "Brain-computer interface technology: a review of the first international meeting". In: *IEEE Transactions on Rehabilitation Engineering* 8.2 (2000), pp. 164–173. DOI: `10.1109/tre.2000.847807`.

[5] Felix Putze et al. "Brain-computer interfaces and augmented/virtual reality". In: *Frontiers in human neuroscience* 14 (2020), p. 144.

[6] Christoph Tremmel, Christian Herff, and Dean J. Krusienski. "EEG Movement Artifact Suppression in Interactive Virtual Reality". In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2019). DOI: `10.1109/embc.2019.8856961`.

[7] Benjamin J. Li et al. "A Public Database of Immersive VR Videos with Corresponding Ratings of Arousal, Valence, and Correlations between Head Movements and Self Report Measures". In: *Frontiers in Psychology* 8 (2017). DOI: `10.3389/fpsyg.2017.02116`.

[8] Soraia M Alarcao and Manuel J Fonseca. "Emotions recognition using EEG signals: A survey". In: *IEEE Transactions on Affective Computing* 10.3 (2017), pp. 374–393.

[9] Beatrice De Gelder. "Why bodies? Twelve reasons for including bodily expressions in affective neuroscience". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364.1535 (2009), pp. 3475–3484.

[10] Kyandoghere Kyamakya et al. *Recent Advances in Nonlinear Dynamics and Synchronization: Theory and Applications*. Vol. 254. Springer, 2009.

[11] Robert Plutchik. "The Nature of Emotions: Clinical Implications". In: *Emotions and Psychopathology* (1988), pp. 1–20. DOI: `10.1007/978-1-4757-1987-1_1`.

[12] James A. Russell. "A circumplex model of affect." In: *Journal of Personality and Social Psychology* 39.6 (1980), pp. 1161–1178. DOI: `10.1037/h0077714`.

[13] Margaret M. Bradley and Peter J. Lang. "Measuring emotion: The self-assessment manikin and the semantic differential". In: *Journal of Behavior Therapy and Experimental Psychiatry* 25.1 (1994), pp. 49–59. DOI: `10.1016/0005-7916(94)90063-9`.

[14] Christoph Tremmel et al. "Estimating Cognitive Workload in an Interactive Virtual Reality Environment Using EEG". In: *Frontiers in Human Neuroscience* 13 (2019). DOI: `10.3389/fnhum.2019.00401`.

[15] You-Yun Lee and Shulan Hsieh. "Classifying Different Emotional States by Means of EEG-Based Functional Connectivity Patterns". In: *PLoS ONE* 9.4 (2014). DOI: `10.1371/journal.pone.0095415`.

[16] Andrea Kübler et al. "Brain–computer communication: Unlocking the locked in." In: *Psychological bulletin* 127.3 (2001), p. 358.

[17] LI Aftanas and SA Golocheikine. "Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution EEG investigation of meditation". In: *Neuroscience letters* 310.1 (2001), pp. 57–60.

[18] Thalıa Fernández et al. "EEG activation patterns during the performance of tasks involving different components of mental calculation". In: *Electroencephalography and clinical neurophysiology* 94.3 (1995), pp. 175–182.

[19] Louise Venables and Stephen H Fairclough. "The influence of performance feedback on goal-setting and mental effort regulation". In: *Motivation and Emotion* 33.1 (2009), pp. 63–74.

[20] Jaime A Pineda. "The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing"". In: *Brain research reviews* 50.1 (2005), pp. 57–68.

[21] Gert Pfurtscheller and Christa Neuper. "Motor imagery and direct brain-computer communication". In: *Proceedings of the IEEE* 89.7 (2001), pp. 1123–1134.

[22] Kwang-Hyuk Lee et al. "Synchronous gamma activity: a review and contribution to an integrative neuroscience model of schizophrenia". In: *Brain Research Reviews* 41.1 (2003), pp. 57–78.

[23] Matthias M Müller et al. "Processing of affective pictures modulates right-hemispheric gamma band EEG activity". In: *Clinical Neurophysiology* 110.11 (1999), pp. 1913–1920.

[24] Li Zhang et al. "Improving mental task classification by adding high frequency band information". In: *Journal of medical systems* 34.1 (2010), pp. 51–60.

[25] Nazmi Sofian Suhaimi, James Mountstephens, and Jason Teo. "EEG-Based Emotion Recognition: A State-of-the-Art Review of Current Trends and Opportunities". In: *Computational Intelligence and Neuroscience* 2020 (2020), pp. 1–19. DOI: 10.1155/2020/8875426.

[26] Nattapong Thammasan et al. "Familiarity effects in EEG-based emotion recognition". In: *Brain informatics* 4.1 (2017), pp. 39–50.

[27] Anne-Marie Brouwer et al. "Estimating workload using EEG spectral power and ERPs in the n-back task". In: *Journal of neural engineering* 9.4 (2012), p. 045008.

[28] Lisa-Marie Vortmann, Felix Kroll, and Felix Putze. "EEG-based classification of internally-and externally-directed attention in an augmented reality paradigm". In: *Frontiers in human neuroscience* 13 (2019), p. 348.

[29] Mariano Alcañiz Raya et al. "The EMMA Project: Emotions as a Determinant of Presence." In: *PsychNology Journal* 1 (Jan. 2003), pp. 141–150.

[30] Nikolaos Malandrakis et al. "A supervised approach to movie emotion tracking". In: June 2011, pp. 2376–2379. DOI: 10.1109/ICASSP.2011.5946961.

[31] Richard W Homan, John Herman, and Phillip Purdy. "Cerebral location of international 10–20 system electrode placement". In: *Electroencephalography and clinical neurophysiology* 66.4 (1987), pp. 376–382.

[32] Riyanarto Sarno et al. "Real-time electroencephalography-based emotion recognition system". In: *Int. Rev. Comput. Softw. IRECOS* 11.5 (2016), pp. 456–465.

[33] Edgar P. Torres et al. "EEG-Based BCI Emotion Recognition: A Survey". In: *Sensors* 20.18 (2020). ISSN: 1424-8220. DOI: 10.3390/s20185083. URL: https://www.mdpi.com/1424-8220/20/18/5083.

[34] Benjamin Blankertz et al. "Optimizing Spatial filters for Robust EEG Single-Trial Analysis". In: *IEEE Signal Processing Magazine* 25.1 (2008), pp. 41–56. DOI: 10.1109/msp.2008.4408441.

[35] Florin Popescu et al. "Single Trial Classification of Motor Imagination Using 6 Dry EEG Electrodes". In: *PLoS ONE* 2.7 (2007). DOI: 10.1371/journal.pone.0000637.

[36] N.J. Hill et al. "Classifying EEG and ECoG signals without subject training for fast BCI implementation: comparison of nonparalyzed and completely paralyzed subjects". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 14.2 (2006), pp. 183–186. DOI: 10.1109/tnsre.2006.875548.

[37] JG Stinstra and MJ Peters. "The volume conductor may act as a temporal filter on the ECG and EEG". In: *Medical and Biological Engineering and Computing* 36.6 (1998), pp. 711–716.

[38] Ian T Jolliffe and Jorge Cadima. "Principal component analysis: a review and recent developments". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2065 (2016), p. 20150202.

[39] Te-Won Lee et al. "Blind source separation of more sources than mixtures using overcomplete representations". In: *IEEE signal processing letters* 6.4 (1999), pp. 87–90.

[40] Aapo Hyvarinen, J Karhunen, and E Oja. *Independent component analysis and blind source separation*. 2001.

[41] Filippo Zappasodi et al. "Age-Related Changes in Electroencephalographic Signal Complexity". In: *PLOS ONE* 10.11 (2015). DOI: `10.1371/journal.pone.0141995`.

[42] C Lithari et al. "Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions". In: *Brain topography* 23.1 (2010), pp. 27–40.

[43] Monson H Hayes. *Statistical digital signal processing and modeling*. John Wiley & Sons, 2009.

[44] Jacob Benesty et al. "Pearson correlation coefficient". In: *Noise reduction in speech processing*. Springer, 2009, pp. 1–4.

[45] Patrick Schober, Christa Boer, and Lothar A. Schwarte. "Correlation Coefficients". In: *Anesthesia amp; Analgesia* 126.5 (2018), pp. 1763–1768. DOI: `10.1213/ane.0000000000002864`.

[46] Laura Martínez-Tejada et al. "Exploring EEG Characteristics to Identify Emotional Reactions under Videogame Scenarios". In: *Brain Sciences* 11.3 (2021), p. 378. DOI: `10.3390/brainsci11030378`.

[47] Diah Risqiwati et al. "Feature Selection for EEG-Based Fatigue Analysis Using Pearson Correlation". In: *2020 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. 2020, pp. 164–169. DOI: `10.1109/ISITIA49792.2020.9163760`.

[48] Michael R Osborne. "Fisher's method of scoring". In: *International Statistical Review/Revue Internationale de Statistique* (1992), pp. 99–117.

[49]   Dennis J McFarland and Jonathan R Wolpaw. "Sensorimotor rhythm-based brain-computer interface (BCI): feature selection by regression improves performance". In: *IEEE transactions on neural systems and rehabilitation engineering* 13.3 (2005), pp. 372–379.

[50]   Isabelle Guyon et al. "A scaling law for the validation-set training-set size ratio". In: *AT&T Bell Laboratories* 1.11 (1997).

[51]   Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. "Brain computer interfaces, a review". In: *sensors* 12.2 (2012), pp. 1211–1279.

[52]   Reinhold Scherer et al. "An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate". In: *IEEE Transactions on Biomedical Engineering* 51.6 (2004), pp. 979–984.

[53]   Christopher JC Burges. "A tutorial on support vector machines for pattern recognition". In: *Data mining and knowledge discovery* 2.2 (1998), pp. 121–167.

[54]   Deon Garrett et al. "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification". In: *IEEE Transactions on neural systems and rehabilitation engineering* 11.2 (2003), pp. 141–144.

[55]   Thomas M Cover. "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition". In: *IEEE transactions on electronic computers* 3 (1965), pp. 326–334.

[56]   Haihong Zhang, Cuntai Guan, and Chuanchu Wang. "Asynchronous P300-based brain–computer interfaces: A computational approach with statistical models". In: *IEEE Transactions on Biomedical Engineering* 55.6 (2008), pp. 1754–1763.

[57] Jie Dou et al. "Evaluating GIS-based multiple statistical models and data mining for earthquake and rainfall-induced landslide susceptibility using the LiDAR DEM". In: *Remote Sensing* 11.6 (2019), p. 638.

[58] K. Tanaka et al. "Stepwise Feature Selection by Cross Validation for EEG-based Brain Computer Interface". In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. 2006, pp. 4672–4677. DOI: `10.1109/IJCNN.2006.247119`.

[59] *VIVE United States: Next-level VR Headsets and Apps*. URL: `https://www.vive.com/us/`.

[60] *Unity Real-time Development Platform*. `https://unity.com/`.

[61] Diederick C Niehorster, Li Li, and Markus Lappe. "The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research". In: *i-Perception* 8.3 (2017), p. 2041669517708205.

[62] Miguel Borges et al. "HTC Vive: Analysis and Accuracy Improvement". In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2018, pp. 2610–2615. DOI: `10.1109/IROS.2018.8593707`.

[63] Xiao Jiang, Gui-Bin Bian, and Zean Tian. "Removal of artifacts from EEG signals: a review". In: *Sensors* 19.5 (2019), p. 987.

[64] Vojkan Mihajlović, Shrishail Patki, and Bernard Grundlehner. "The impact of head movements on EEG and contact impedance: An adaptive filtering solution for motion artifact reduction". In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2014, pp. 5064–5067.

[65] Ian Daly et al. "On the Automated Removal of Artifacts Related to Head Movement From the EEG". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21.3 (2013), pp. 427–434. DOI: `10.1109/TNSRE.2013.2254724`.

[66] A. Belouchrani et al. "A blind source separation technique using second-order statistics". In: *IEEE Transactions on Signal Processing* 45.2 (1997), pp. 434–444. DOI: `10.1109/78.554307`.

[67] MATLAB. *version 9.7.0.1319299 (R2019b)*. Natick, Massachusetts: The MathWorks Inc., 2019.

[68] Sander Koelstra et al. "Deap: A database for emotion analysis; using physiological signals". In: *IEEE transactions on affective computing* 3.1 (2011), pp. 18–31.

[69] Nicole A. Lazar et al. "Combining brains: A survey of methods for statistical pooling of information". English (US). In: *NeuroImage* 16.2 (June 2002), pp. 538–550. ISSN: 1053-8119. DOI: `10.1006/nimg.2002.1107`.

[70] Wolfgang Klimesch, Paul Sauseng, and Simon Hanslmayr. "EEG alpha oscillations: the inhibition–timing hypothesis". In: *Brain research reviews* 53.1 (2007), pp. 63–88.

[71] Harry W Cole and William J Ray. "EEG correlates of emotional tasks related to attentional demands". In: *International Journal of Psychophysiology* 3.1 (1985), pp. 33–41.

[72] Irina I Goncharova et al. "EMG contamination of EEG: spectral and topographical characteristics". In: *Clinical neurophysiology* 114.9 (2003), pp. 1580–1593.