AFFECTIVE VISUALIZATION IN VIRTUAL REALITY SYSTEMS

vorgelegt von
Andres Pinilla Palacios
ORCID: 0000-0002-0812-7896

an der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
und der
University of Technology Sydney, Australia
(im Rahmen des Doppel-Promotionsabkommens)
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr. rer. nat. -

genehmigte Dissertation

Promotionsausschuss:
Vorsitzender: Prof. Dr.-Ing. Sebastian Möller (TU Berlin)
Gutachter: Univ.-Prof. Dipl.-Ing. Dr.Helmut Hlavacs (Universität Wien)
Gutachter: Dr. Yusuf Pisan (University of Washington)
Tag der wissenschaftlichen Aussprache: 28. Februar 2022
an der Technischen Universität Berlin

Berlin 2022

UNIVERSITY
OF TECHNOLOGY
SYDNEY

AFFECTIVE VISUALIZATION IN VIRTUAL REALITY SYSTEMS

by Andres Pinilla Palacios

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Science

under the supervision of
Jaime Garcia (UTS)
Sebastian Möller (TU Berlin)
Jan-Niklas Voigt-Antons (TU Berlin)
William Raffe (UTS)

University of Technology Sydney
Faculty of Engineering and Information Technology
And
Technische Universität Berlin
Faculty IV - Electrical Engineering and Informatics

February 2022

Certificate of Original Authorship

I, Andres Pinilla Palacios declare that this thesis is submitted in fulfilment of the requirements for the award of doctorate degree, in the Faculty of Engineering and IT at the University of Technology Sydney, and the Faculty of Computer Science and Electrical Engineering at the Technische Universität Berlin.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree at any other academic institution except as fully acknowledged within the text. This thesis is the result of a Collaborative Doctoral Research Degree program with the Technische Universität Berlin.

This research is supported by the Australian Government Research Training Program and by the strategic partnership between the both universities.

Signature: Andres Pinilla Palacios

Digitally signed by Andres Pinilla Palacios
Date: 2022.04.28 11:07:32 +02'00'

Date: April 25, 2022

Zusammenfassung

Eine Reihe von Forschungsarbeiten auf dem Gebiet der Mensch-Computer-Interaktion (HCI) legt nahe, dass es möglich ist, durch die Analyse der elektrophysiologischen Reaktionen der Benutzer in Echtzeit auf einige Merkmale ihrer mentalen Verfassung zu schließen. Es ist jedoch unklar, wie die aus den elektrophysiologischen Signalen gewonnenen Informationen genutzt werden können, um die Stimuli innerhalb einer virtuellen Umgebung entsprechend dem affektiven Zustand des Benutzers anzupassen. Daher ist das Hauptziel dieses Forschungsprojekts zu verstehen, wie ein VR-System entwickelt werden kann, das sich automatisch an den affektiven Zustand des Benutzers anpasst. Es wird eine Referenzimplementierung einer Neurofeedback-VR-Erfahrung zum Training der affektiven Selbstregulierung vorgeschlagen. Diese Erfahrung zielt darauf ab, die Fähigkeit der Benutzer zu trainieren, ihre affektiven Zustände freiwillig zu regulieren. Die wichtigsten Beiträge sind (1) die Entwicklung einer Technik zur Erkennung affektiver Zustände bei VR-Nutzern nahezu in Echtzeit; (2) eine virtuelle Umgebung zur visuellen Darstellung affektiver Zustände; und (3) die Implementierung der Technik zur Erkennung von Affekten und der virtuellen Umgebung für die Entwicklung einer Neurofeedback-VR-Erfahrung.

Abstract

A cluster of research in Human-Computer Interaction (HCI) suggests that it is possible to infer some characteristics of users' mental states by analyzing their electrophysiological responses in real-time. However, it is unclear how to use the information extracted from electrophysiological signals to adjust the stimuli inside a virtual environment according to the user's affective state. Therefore, this research project's main objective is to understand how to develop a VR system that adapts automatically to the user's affective state. A reference implementation of a neurofeedback VR experience for training affective self-regulation is proposed. This experience aims to train the ability of users to regulate their affective states voluntarily. The main contributions are (1) the development of a technique for near real-time detection of affective states in VR users; (2) a virtual environment for visual representation of affective states; and (3) the implementation of the affect detection technique and the virtual environment for the development of a neurofeedback VR experience.

Acknowledgements

This thesis was possible thanks to a joint doctoral program established between TU Berlin and the University of Technology Sydney. Participating in this program gave me the freedom to investigate the topic I was interested in. At the same time, it provided me the privilege of living in two countries during my studies. Moreover, gave me the opportunity of meeting some of the smartest people I have met. I am sincerely thankful to all the people in both institutions who made this possible.

Thanks to Sebastian Möller for his generosity, constant support, and for giving me the privilege of being part of his team for a short time.

Thanks to Jan-Niklas Voigt-Antons for opening the doors for me in this program. Moreover, thanks for trusting in me without knowing me.

Thanks to Jaime Garcia for his constant support, kindness, and for being always there.

Thanks to William Raffe for his accurate and timely advice.

Thanks to Irene Hube-Achter and Yasmin Hillebrenner for their help in solving all the unexpected issues during this process.

Thanks to Ana Paula Ruiz Ochoa for her support at UTS and her warm welcome in Sydney.

Thanks to Evelina Skurski, Johannes Moes, and Jing Zhao for their help establishing the collaboration between both universities.

Thanks to Gail Kenning for her advice and for participating in my Candidate Assessment sessions.

Thanks to Nick Busietta for providing me access to the PsyDocs of LiminalVR. Those documents were helpful for the development of the virtual environment proposed in this thesis.

Thanks to my colleagues at the Quality and Usability Lab for their support and all the coffees that we had. Especially, thanks to Kerstin Pieper and Stefan Uhrig for sharing their knowledge about electrophysiology, and to Danish Ali for his help with VR development.

Thanks to Avinash Singh and professor CT Lin for their support and opening the doors for me at their lab.

Thanks to Prith and Abhi for helping to find participants for my experiments. Recruiting people would have been much harder without your help.

Thanks to Ricardo Tamayo for his generous and selfless advice.

Thanks to Tilo, Sonia, Fiza, Mauricio, Niveditha, Pratik, and Diego. I was lucky to cross paths with you.

Thanks to my dear Elena for your constant support and advice. This process would have been much more difficult without you.

Thanks to my sister, Natalia, and my brother, Jorge, because you have been always around when needed. And most of all, thanks to my parents, Sofía and Humberto, because you have given me everything.

Acronyms

- Action Unit (AU)
- Artifact Subspace Reconstruction (ASR)
- Brain–Computer Interfaces (BCIs)
- Electrocardiography (ECG)
- Electroencephalography (EEG)
- Electromyography (EMG)
- Evaluative Space Model (ESM)
- Facial Action Coding System (FACS)
- Head–Mounted Display (HMD)
- Heart Rate Variability (HRV)
- Human Computer Interaction (HCI)
- Independent Component Analysis (ICA)
- International Affective Pictures System (IAPS)
- Long Short-Term Memory Recurrent Neural Networks (LSTMRNN)
- Low Frequency / High Frequency ratio (LF/HF ratio)
- Ortony, Clore & Collins theory of emotions (OCC theory)
- Pick a Mood (PAM)
- Positive and Negative Affect Schedule (PANAS)
- RR-Intervals (RRI)
- Root Mean Square of Successive Differences (RMSSD)
- Self-Assessment Manikin (SAM)
- Standard Deviation of NN intervals (SDNN)
- Virtual Reality (VR)

Table of Contents

1 Introduction

The prevalence of mental disorders is increasing worldwide. Between 2007 and 2017, mental health and substance abuse disorders increased by 13% [1]. Some of the most common mental disorders (depression, bipolar disorder, schizophrenia, and dementia) affected around 379 million people in 2019 [2]. It is expected that the number of people suffering from mental disorders will increase during the following years due to demographic factors. For example, it is expected that the elderly population will be doubled by 2050 [3]. Given that older adults are more prone to suffering from dementia, it is expected that the number of people affected by this disorder will tend to increase.

At the same time, the World Health Organization (WHO) pointed out the lack of access to mental healthcare across the globe and proposed a plan to alleviate this issue [4]. Lack of access to mental healthcare might aggravate existing mental disorders, with particularly severe consequences for vulnerable population (e.g., older adults, victims of armed conflicts, and people who live in extreme poverty). The problem worsened after the COVID-19 pandemic because the demand for mental healthcare services increased during the crisis [5, 6]. Countries with well-established healthcare systems, such as Germany and Australia, could not respond to this increased demand. Partly, because there are not enough qualified professionals [7, 8].

This challenge could be mitigated by developing technologies that provide mental healthcare without the constant supervision of a specialist. These technologies could lead to cost-effective solutions, while fostering the development of innovative tools for mental healthcare. Emotions are associated with some of the most common mental disorders. For example, depression and anxiety are associated with negative emotions [9], while psychopathy is associated with impaired emotional empathy [10]. Therefore, technologies for mental healthcare could be improved with systems that automatically analyze patients' emotional states.

Existing technologies for mental healthcare usually rely on a User Interface (UI) that allows interacting with a computer through a 2D display. However, recent advances in Virtual Reality (VR) technologies lead to new possibilities for interaction between humans and computers. Moreover, VR provides the possibility of entering into digital worlds that would be impossible to build in the physical reality. Those digital worlds can be used for the development of mental health therapies. Virtual Reality Exposure Therapy (VRET) [11], for instance, allows creating safe spaces where patients that suffer phobias are exposed systematically to the stimulus that triggers the phobic reaction. Eventually, the patient is desensitized and learns to handle the phobia.

There are at least five advantages in using VR systems for mental healthcare therapies. Firstly, VR systems are more effective than 2D displays at blocking out stimuli from the physical world. Thus, VR systems are more immersive [12]. Secondly, interactive experiences delivered through VR systems tend to produce more intense emotional responses than experiences delivered through 2D displays [13, 14]. Thirdly, VR systems provide high degree of control over the stimuli, allowing to expose patients to scenarios that could be dangerous in the physical world [11]. Fourthly, VR systems allow creating synthetic worlds with high levels of realism, or even create virtual worlds with different levels of realism [15]. Fifthly, VR systems can be used to build games for health [16, 17, 18], fostering patients' engagement with non-supervised therapeutic tools.

Usually, a VR system is controlled by the user through keys pressed in a controller. However, previous studies indicate that it is possible to interpret users' affective states as commands [19, 20]. Those commands can be used to adjust the VR content automatically [21]. In general, VR content consists of stimuli that can be delivered through different sensory channels, such as the visual, auditory, haptic or olfactory modalities. Previous research indicates that humans tend to prioritize the processing of visual stimuli over other types of stimuli [22]. Therefore, this thesis is focused on visual stimuli.

1.1 Aim, Research Questions and Scope

1.1.1 Aim

This research project aims to understand how to build a VR system that automatically adapts its visual stimuli according to the user's affective state. A reference implementation of a neurofeedback VR system for training affective self-regulation was built for this purpose. In this system, users recall autobiographical memories that are associated with their affective states. The brain activity of the user is monitored using electroencephalography (EEG) sensors. The EEG signals are processed in near real-time to estimate the affective state of the user. Then, visual feedback about the affective states evoked by the user is provided inside the virtual environment.

1.1.2 Research questions

This thesis addresses three main questions:

1. How to build a virtual environment for affective visualization? Developing a virtual environment for affective visualization would be helpful for providing visual feedback to VR users about their affective states. There are countless aesthetic possibilities for creating visualizations, and individual differences between users lead to endless possible interpretations of a single visualization. However, previous studies suggest an association between visual cues (e.g., color, shape, texture, motion patterns) and affective states (see section 2.2). Analyzing those previous studies might allow identifying general guidelines for designing affective visualizations (see Table 2.1). Moreover, those guidelines can be used to develop the virtual environment used in the proposed reference implementation (see chapter 3).

2. How to develop a technique for near real-time affect detection that can be implemented in VR systems? Ideally, a technique for affect detection in VR should estimate the users' affective states in real-time or near real-time. At the same time, the technique should use non-intrusive sensors to minimise any potential risks for the user. Previous studies indicate that both requirements can be achieved using electroencephalography (EEG) signals (see section 2.3). Yet, there is no clarity about the precise signal processing steps required to infer users' affective states from EEG signals.

3. How to integrate the virtual environment with the affect detection technique? The proposed reference implementation requires integrating the virtual environment (first research question) with the affect detection technique (second research question). This would allow building a system to automatically analyze users' affective states. Affective visual feedback would be provided inside the virtual environment. This approach can be used to design interactive experiences where the desired conduct (e.g., evoking positive affective states) is reinforced using reward systems (e.g., game score). Integrating the virtual environment with the affect detection technique requires building an interface for connecting both components. Previous studies indicate that it is possible to build Brain-Computer Interfaces (BCIs) for that purpose (see section 2.4). However, there is not enough research about building a BCI for affect detection that can be integrated with a VR system.

1.1.3 Out of scope

This thesis does not attempt to evaluate therapeutic methods. Instead, it aims to understand how to develop a system that can be used in VR therapies. The clinical evaluation of therapeutic methods based on VR systems is out of the scope of this manuscript. Therefore, no clinical studies are conducted. A clinical evaluation would require the involvement of patients with a history of mental disorders. In this research project, participants were adults without a history of mental disorders.

Given that the proposed reference implementation aims to infer users' affective states from their brain activity, it requires building an interface between the users' brain and the VR system. Therefore, a Brain-Computer Interface (BCI) is developed. Yet, the proposed BCI has not been tested in an experiment with human participants. An experiment was planned to test the system, but it was cancelled due to COVID-19 restrictions. Still, the proposed reference implementation might be useful for research and further development.

Procedural Content Generation (PCG) and its potential for creating personalized affective visualizations is discussed. However, PCG is not implemented in the proposed reference implementation. Instead, basic if/then rules are used for controlling the stimuli in the virtual environment.

1.2 Contributions

The three main contributions of this thesis are:

1. Development and validation of four virtual environments for affect visualization.

2. Develop a technique for near real-time affect detection. The technique was tested in an offline setting, emulating the process that would be conducted in a real-time implementation.

3. A reference implementation of a neurofeedback VR system that can be used for research purposes and further improvement.

1.3 Research phases

This research project was undertaken in four phases:

Phase 1: Literature review

- Analyze concepts and theories related to affect and emotion.
- Summarise previous studies related to the affective interpretation of audio-visual cues.
- Identify available methods for assessment of affective states.
- Explore existing technologies for the development of adaptive VR systems.

Phase 2: Development of VR content.

- Identify guidelines for the design of affective visualizations.
- Design and build the virtual environments.
- Test the virtual environments in an experiment with human participants.

Phase 3: Development of affect detection technique.

- Analyze how to infer affective states in VR users.
- Propose a feature extraction process and test it with the preprocessed version of the DEAP dataset [23].
- Incorporate the feature extraction process into a technique for near real-time affect detection.
- Test the proposed affect detection technique with a standard dataset for emotion research [23], as well as with data collected during a pilot study in VR.

Phase 4: Development of reference implementation.

- Define system architecture.
- Integrate the two main components of the prototype: the VR content module and the affect detection module.
- Test the reference implementation in an experiment with human participants (not completed because of COVID-19 restrictions).

1.4 Thesis Structure

Chapter 1 introduces the research problem, research questions, contributions, and scope of the thesis. Chapter 2 reviews relevant literature, integrating previous studies in psychology, electrophysiology, audio-visual design, and BCIs. Chapter 3 describes the development and validation of four virtual environments for affect visualization. Chapter 4 proposes an affect detection technique based on EEG signals. Chapter 5 presents Sentui, a reference implementation of a VR neurofeedback system developed based on the results presented in chapters 3 and 4.

Previous studies reviewed in section 2.2 allowed to propose guidelines for the design of stimuli for visual representation of affective states. Based on those studies, four virtual environments were designed. The design guidelines presented in section 2.2 (see Table 2.1) were used to develop the VR content that was tested in the experiment presented in chapter 3. This chapter responds the first research question of this thesis.

Affect assessment methods were reviewed in section 2.3 and contributed to understand how to develop the affect detection technique proposed in chapter 4. Therefore, this chapter responds the second research question of this thesis. Chapter 4 consists of two main sections. The first section consist of the validation of the feature extraction process, using the preprocessed version of the Database for Emotion Analysis using Physiological Signals (DEAP dataset) [23]. The second section consists of a pilot study that aims to analyze the feasibility of implementing the proposed affect detection technique in a VR system, proposing a a method suitable for near real-time affect detection.

A reference implementation of a neurofeedback VR system is presented in chapter 5. This reference implementation consists of two modules: a VR content module and an affect detection module. The VR content module was developed using the virtual environments presented in chapter 3, while the affect detection module was developed using the affect detection method proposed in chapter 4. The integration between both modules allows building a neurofeedback VR system for training affective self-regulation. Therefore, chapter 5 responds the third research question of this thesis. Finally, chapter 6 consists of a general discussion, followed by an analysis of the limitations of this work and questions that could be addressed in future research projects.

1.5 Publications

Parts of this thesis were previously published as peer-reviewed articles and preprints in three manuscripts:

1. A. Pinilla, J. Garcia, W. Raffe, J.-N. Voigt-Antons, R. P. Spang, and S. Möller, "Affective Visualization in Virtual Reality: An Integrative Review," Front. Virtual Real., vol. 2, p. 630731, 2021, doi: 10.3389/frvir.2021.630731.

Author Contributions: All authors contributed to the study conception and analysis. Jaime Garcia and William Raffe contributed with analysis of data related to gaming and Virtual Reality. Jan-Niklas Voigt-Antons contributed to the analysis of data related to psychology and electrophysiology. Robert Spang contributed with the redaction of the manuscript and with data related to electrocardiography and eye-tracking. Sebastian Möller contributed to the analysis of data related to sound design and Machine Learning. Andres Pinilla conducted the literature search, data analysis and wrote the first draft. All authors commented on previous versions of the manuscript.

2. Pinilla, A., Garcia, J., Raffe, W., Voigt-Antons, J.N., Möller S. (2021). Visual representation of emotions in Virtual Reality. PsyArXiv. 10.31234/osf.io/9jguh

Author Contributions: All authors contributed to the study conception and data analysis. Jaime Garcia and William Raffe contributed with the development of the virtual environment. Jan-Niklas Voigt-Antons and Sebastian Möller contributed with the experimental design. Andres Pinilla conducted the study, developed the virtual environments and wrote the first draft. All authors commented on previous versions of the manuscript.

3. Pinilla, A., Voigt-Antons, J.N., Garcia, J., Raffe, W., Möller S. (2021a). Affect detection in Virtual Reality: An exploratory study. Unpublished manuscript, Quality and Usability Lab, Technische Universität Berlin, Berlin, Germany. UTS Games Studio, University of Technology Sydney, Sydney, Australia.

Author Contributions: Jan-Niklas Voigt-Antons and Sebastian Möller contributed with the study conception and experimental design. Jaime Garcia and William Raffe proof-read the manuscript. Andres Pinilla conducted the study, developed the virtual environments and wrote the first draft. All authors commented on previous versions of the manuscript.

2 Background

This chapter is based on a review paper [24]. Some fragments of that paper were written by the coauthors. Those fragments were excluded from this chapter.

Virtual Reality (VR) systems offer endless possibilities for the development of interactive experiences. They are used for the development of tools in diverse areas such as rehabilitation therapy [25], exergames [26, 27], and robotics [28]. Their potential is particularly promising when combined with technological advances in Affective Computing, allowing to interpret users' affective states as computer commands [29, 30, 31, 32] and adapt the content of a virtual environment accordingly [33].

Traditional psychological tasks for the treatment and diagnosis of mental disorders can be replaced by VR systems [34, 35, 36, 18, 37]. These new tools are less time-consuming and provide more realistic environments than traditional tools for mental healthcare [38]. Furthermore, VR might be helpful for at least two types of therapy: exposure therapy and biofeedback therapy. Exposure therapy is commonly used to treat anxiety disorders caused by phobias. Patients are systematically exposed to the stimuli that trigger the phobia in a controlled environment. VR is useful for exposure therapy because it allows delivering realistic experiences while providing control over the stimuli. Previous research suggests that exposure therapy in VR might be effective for the treatment of a least three types of phobias: social phobia [11], claustrophobia [39], and spider-phobia [40].

Similarly, Cavazza et al. [41] developed an interactive experience to enhance empathy using neurofeedback. The participant interacted with a fictional doctor that was going through a difficult situation. Simultaneously, the participant's EEG signals were analyzed to estimate the affective response towards the doctor. If the system detected a positive affective response in the player, the storyline would change positively (the doctor would struggle less), reinforcing supportive and empathetic behavior in the player.

Patients who lack affective self-regulation could benefit from VR biofeedback therapy to train affective self-control, fostering mood regulation [42, 43]. Li et al. [44] conducted an experiment where twenty-three participants' brain activity was analyzed in real-time using functional Magnetic Brain Imaging (fMRI). The researchers asked participants to evoke a happy or sad memory and provided feedback about their affective state. The feedback was provided with a bar on a screen. The bar's level increased when the fMRI data indicated that the participant successfully evoked the happy or sad memory. Results suggest that providing visual feedback allowed participants to learn how to modulate their neural activity. However, it is not clear how to implement this finding in a therapeutic application that can be accessible for a large population because 1) fMRI is an expensive technology that is not accessible for most people, 2) participants must remain motionless for long periods during fMRI recordings, and 3) an ideal therapeutic application should consist of engaging content that motivates users to use the system. These challenges could be solved by measuring brain activity with a less expensive and more portable method than fMRI, such as electroencephalography (EEG). The visual feedback could be provided using game-like elements in VR.

Developing an affective visualization tool in VR would require at least two components: (1) A set of VR stimuli with affective content whose properties can be adjusted in real-time, and (2) a technique to continuously assess affective states without interrupting the VR experience. This literature review was elaborated to understand how to develop those components. Both requirements are addressed in three subsections. Firstly, theories related to emotions and affect are presented. Secondly, findings related to visual and sound cues that are associated with affective responses are analyzed. And thirdly, some of the most common methods for detecting affective states are summarized.

2.1 Theoretical models of emotion and affect

The terms emotion and affect are often used interchangeably in the literature, but they are not the same. There is not a general agreement about how to define these concepts. In this thesis, emotions are defined as mental states that coordinate the operation of cognitive processes. This definition is based on the assumption that the human mind is designed as a computational system that consists of a series of information-processing programs [45]. Emotions are a particular type of program that coordinates other programs' operation [46]. Affect is defined as the cognitive representation of the bodily changes that come with emotions [47, 48]. Neither emotion nor affect can be directly observed or measured. However, the construct of affect is directly associated with physiological states. Therefore, it is reasonable to use electrophysiological signals to infer affective states, which in turn might allow inferring some characteristics of emotional states.

2.1.1 Emotion theories

The Ortony, Clore and Collins (OCC) theory of emotions [49] has been widely used in the field of computer science to model users' emotional responses, e.g. [50, 51]. This theory describes emotions in terms of twenty-two categories and assumes a clear distinction between each category. This approach is compatible with existing emotion recognition algorithms because these are usually based on categorizing emotional responses [52, 53]. According to the OCC theory [49], the first step in an emotional response is the perception of the situation. Then the situation is evaluated (appraisal), and finally, the emotional response emerges. However, this theory does not consider the physiological changes associated with emotions.

Similarly, Robert Plutchnik proposed a structural model of emotions [54], commonly known as Plutchnik's wheel of emotions. This model consists of eight primary states: ecstasy, adoration, terror, amazement, grief, loathing, rage, and vigilance. According to Pluthnik's theory, any emotion can be described as a combination of a subset of those primary states. Here emotions are defined as a sequence of reactions towards a stimulus. This sequence includes a cognitive evaluation of the stimuli (appraisal), feelings (subjective experience of the emotion), autonomic neural activity, and behavioral responses.

There are at least three other major emotion theories: the James-Lang Theory [55], the Cannon-Bard Theory [56, 57], and the Schachter-Singer Theory [58]. According to Shiota and Kalat [59], these theories have in common the assumption that emotional responses have four components but differ in the order those components take place during an emotional response. The components are:

- **Appraisal:** The cognitive, rationalized evaluation of the context where the emotional response is produced.
- **Feeling:** The subjective, momentary experience of the emotion.
- **Physiological change:** The bodily changes produced by the emotional response.
- **Behavior:** The observable conduct that comes with the emotion.

According to the James-Lange theory [55], the first step in an emotional response is the cognitive evaluation of the situation. Then, physiological changes are produced in the body, at the same time that a behavioral response is produced. Lastly, feelings take place.

The Cannon-Bard Theory [56, 57] proposes that all the elements of an emotional response are independent of each other, and there is no particular order in which they occur. This theory is not compatible with strong evidence indicating that emotional stimuli tend to trigger automatic changes in the body, e.g. [60, 61, 62]. Overall, those previous studies suggest interdependence between physiological states and emotions.

According to the Schachter and Singer theory [58], physiological changes occur first. Then the user tries to find an explanation in the environment for those physiological changes. Depending on the explanation found, a cognitive label is assigned to the bodily changes perceived. Therefore, the physiological changes indicate the intensity of the emotional experience, but cognitive factors determine the emotion's valence (pleasant vs. unpleasant).

2.1.2 Theoretical models of affect

Theoretical models of affect can be classified into two major groups: discrete and dimensional models. Discrete models are based on a categorical division of affective responses, while dimensional models represent affect as an array of continuous variables. Both types of models are commonly used in Affective Computing to build affect recognition models, e.g. [32, 29, 30].

In broad terms, discrete models propose the existence of a few primary states, such as happiness, sadness, and anger. Affective responses are a combination of a subset of those fundamental states. A prominent example of the influence of discrete models in psychology research can be found in an experiment conducted by Ekman and Friesen [63] in New Guinea. In this experiment, stories with emotional content were told to 153 participants. One hundred thirty of them (84.97%) had no previous contact with western culture. After each story had been told, participants saw a series of pictures of facial expressions and were asked to choose the more coherent face with the story. Interestingly, participants associated similar facial expressions to the same stories, regardless of their cultural background. Based on this evidence, it was proposed that at least six facial expressions are universal (i.e., they are not affected by culture): happiness, anger, sadness, disgust, surprise, and fear. These results are consistent with earlier contributions from Charles Darwin, who pointed out the existence of activation patterns in facial muscles that are associated with affective states [64, 63].

Dimensional models have their roots in the early contributions of Wilhelm Wundt, who proposed that affective responses have three dimensions: valence (pleasant–unpleasant), arousal (arousing–subduing), and intensity (strain–relaxation) [47]. On this basis, the Circumplex Model of Affect (CMA) [65] was developed, representing affect in a two-dimensional space, where valence and arousal are equivalent to the x-axis and y-axis, respectively.

Other authors have proposed the Evaluative Space Model (ESM) [66], which has three dimensions: Negativity in the x-axis, Positivity in the y-axis, and Net predisposition (to withdraw or approach a stimulus) in the z-axis. Unlike the CMA [65], the ESM [66] contemplates the existence of affective responses with simultaneous negative and positive activation ("bitter-sweet" affective states). For example, while playing a terror video game, the user might feel fear, and at the same time, might feel excited because is aware that there is not a real danger. An analysis of dimensional models of affect can be found in Mattek et al. [67].

The ESM proposes the existence of the negativity bias and the positivity offset. The negativity bias implies that negative activation produces more changes in the motivation to withdraw or approach stimuli than positive activation. Evidence supporting the existence of the negativity bias indicates that negative stimuli tend to produce more salient behaviors than positive stimuli [68], and negative stimuli tend to be associated with higher arousal than positive stimuli [69]. The negativity bias suggests that terror video games should trigger higher arousal than video games associated with positive affective states. However, a recent study indicates that the arousal triggered by terror video games is slightly lower than the arousal triggered by video games associated with positive affective states [70].

The positivity offset implies a slight positive motivation to approach unknown stimuli in a neutral environment. This mechanism has been associated with humans' natural tendency to explore new, unthreatening environments, even when that behavior is not associated with a reward [66]. Further research about the positivity offset could help understand how to motivate VR users to explore virtual environments. For example, to stimulate engagement of players with VR games.

2.2 Visual and sound cues

Building a virtual environment for affective visualization requires content that any user can associate with a wide range of affective states, regardless of cultural differences or personal preferences. Therefore, this section presents studies indicating an association between some characteristics of graphical elements and affective states. The aim is to provide general guidelines for designing visual and sound elements when creating visual representations of affective states, rather than trying to define a set of rigid rules.

2.2.1 Visual cues

Rounded objects are associated with higher valence and lower arousal than sharp objects [71]. At the same time, rounded lines are perceived as more attractive than straight or angular lines [72, 73]. Given that attractiveness is associated with positive affective states, rounded lines are likely associated with positive valence. Additional studies suggest that visual complexity plays a role in the likability of objects. People tend to prefer extremely simple or extremely complex objects [74]. Given that likability tends to trigger positive valence [75], an intermediate level of complexity is more likely associated with negative valence.

A cross-cultural study showed that the most critical factors in the affective meaning of color are brightness and saturation, while hue has a secondary role [76]. These results are consistent with evidence reported in Valdez and Mehrabian [77] but contrast with recent studies indicating that hue has a significant role in the affective state associated with a color palette [78]. Additional evidence suggests that blue, green, and purple are among the most pleasant hues, while yellow is among the most unpleasant. Green-yellow, blue-green, and green are the most arousing, while purple-blue and yellow-red are among the least arousing [79]. Similarly, it has been found that the most pleasant colors are those with higher saturation and brightness [80, 81]. However, other studies suggest that there are not universal associations between colors and affective states. People tend to like colors associated with objects they like and dislike colors associated with objects they dislike [79]. Additional evidence indicates that color associations change according to the context where colors are used [82]. Yet, it is possible to establish color palettes that allow to communicate affective states. For example, bright, unsaturated colors are more suitable to communicate calm, while dark, red colors are more suitable to communicate disturbance [78].

Textures may influence the affective meaning of color [83, 84]. This has been demonstrated by pairing colors with computer-generated textures and asking participants to rate the color-texture pairs using four scales: Warm-Cool, Masculine-Feminine, Hard-Soft, and Heavy-Light. Results suggest that texture significantly influences the evaluation on the Hard-Soft scale and has a minor impact on the other scales. However, this evidence does not allow to identify associations between particular texture patterns and affective responses.

Non-static visual elements have other visual properties besides color, shape, and texture. Some of these additional properties are speed, motion shape, direction, and path curvature. Fast-moving objects are associated with higher arousal than slow-moving objects [85, 86]. However, there are contradictory findings regarding the type of valence associated with speed. One study suggests that fast movements are related to positive affective states [86], while another study indicates the opposite [85].

Linear motion with straight paths is associated with low arousal and positive valence [85]. Jerky paths are associated with higher arousal than straight paths in linear motion [87, 85]. However, the curvature of the path has no incidence in affective associations when applied to spherical or radial motion [85]. Inward movements are related to more positive affective states than outward movements [85]. Downwards-right motion tends to be linked with positive states, while upwards-left motion tends to be associated with negative states [87]. In general, angular paths are related to more negative affective states than linear paths [87]. Likewise, spherical motion patterns tend to be associated with higher arousal than linear motion patterns.

2.2.2 Sound cues

Previous research indicates that the location of a sound source influences the affective states associated with that sound. When the user cannot see where the object is (outside of the field of view), it is usually associated with higher arousal than when the user can see it (inside the field of view) [88, 89]. Similarly, sounds located further away in the space are related to less arousing responses [90]. The perception of an approaching sound is associated with more arousing responses than the perception of it moving away [91]. These phenomena are likely to be linked to mechanisms enforced by evolution [46]. Our primitive ancestors had more chances to survive if they were aware of the most potentially dangerous objects, such as those they could not see, were closer to them, or were approaching them.

The reverberation of the sound, which is associated with space's size, can influence affective associations [89]. Lower reverberation (smaller rooms) is linked to more pleasant states than higher reverberation (larger rooms). Perhaps, because the primitive human being was better protected from predators in closed spaces, leading to an evolutionary process that favors the activation of attentional resources when people are in open areas.

Other studies indicate that asking people to rate pictures with affective content while listening to the sound of a heartbeat can influence their affective evaluations, as well as their heart rate [90]. In other words, the sound of a heartbeat that is faster than the listener's cardiac rhythm tends to increase the listeners' heart rate. Therefore, playing a fast heartbeat in the background might be an effective way of representing an increase in arousal.

On the other hand, music is pivotal for affective visualization because it can contribute to creating more immersive experiences. However, it is a vast topic that will not be fully covered in this thesis. Yet, it is important to mention that tempo influences music's affective perception [92]. Faster tempo tends to be associated with higher arousal ratings, while slower tempo tends to be associated with lower arousal ratings. To the extent of our knowledge, there is no evidence suggesting that tempo influences valence ratings.

Major and minor chords are associated with positive and negative affective states, respectively [93]. Similarly, dissonant harmonies tend to be strongly associated with anger, and to a lesser extent, with fear [94]. Additional research indicates that it is possible to compose music based on people's affective states [95]. However, it remains an open question whether it is feasible to do it in real-time, based on electrophysiological signals recorded from the user.

2.2.3 Personalized affective visualizations

There might be individual differences in the affective states that each user associates with each audio-visual stimuli. These individual differences could be amplified as a consequence of personal experiences. An ideal system for affective visualization should account for those individual differences, delivering personalized visual representations of affective states, similar to Bermudez i Badia et al., [33].

Semertzidis et al. [21] developed an Augmented Reality (AR) system that automatically creates visual representations of the user's affective states. The visualizations consisted of fractals whose visual properties varied according to the affective state detected in the user. However, the evidence reported by Semertzidis et al. [21] does not allow to establish whether participants perceived that the fractals' graphical properties represented their affective states.

Additional studies indicate that it is possible to use Procedural Content Generation (PCG) to create content dynamically, adjusting it to the user's preferences. This approach is known as

Table 2.1: Summary of audio-visual cues associated with affective states, according to previous studies.

		High Arousal	Low Arousal	High Valence	Low Valence
Static visual cues	Shape	Angular	Rounded	Rounded	Angular
	Lines	N/A	N/A	Rounded	Straight
	Hue	Green-yellow, blue-green, and green	Purple-blue and yellow-red	Blue, green, and purple	Yellow
	Saturation	N/A	N/A	Saturated	Unsaturated
	Brightness	N/A	N/A	Bright	Dark
	Visual complexity	N/A	N/A	Extremely complex or extremely simple	Neither complex nor simple
Non-static visual cues	Speed	Fast	Slow	Some studies suggest that fast motion is associated with positive valence, while others suggest the opposite.	
	Motion shape	Spherical	Linear	N/A	N/A
	Direction	N/A	N/A	Downwards-right, inward	Upwards-left, outward
	Path curvature	Jerky	Straight	Jerky	Straight
Sound cues	Source location	Outside field view	Inside field view	N/A	N/A
	Distance to the sound source	Near	Far	N/A	N/A
	Sound source movement	Approaching	Receding	N/A	N/A
	Heartbeat sound	Fast (above 100 bpm)	Slow (below 60 bpm)	N/A	N/A
Music	Tempo	Fast	Slow	N/A	N/A
	Harmony	N/A	N/A	Major scale	Minor scale

experience-driven procedural content generation (EDPCG) [96, 97]. In broad terms, EDPCG consists of an iterative process where the content is constantly modified based on the user's feedback.

The general functioning of EDPCG is the same as an evolutionary algorithm (EA), which is an optimization process inspired by natural evolution. In a natural environment, the organisms better adapted to their habitat tend to have more reproductive success, hence more likely to pass their genes to the next generation. Similarly, objects can be created programmatically in a virtual environment and tested to identify the most successful ones. The criteria to identify which objects are more successful is based on a previously defined goal. This goal is defined by the developer based on the purpose of the application. During each iteration, the objects that are more successful at reaching the goal are identified. In the following iterations, new sets of objects are created, and the characteristics of the most successful objects tend to remain, whereas the characteristics of the least successful tend to disappear. It is assumed that repeating this process several times allows reaching the optimal parameters required to achieve the goal. For example, the goal could be creating personalized visual representations of positive affective states. If the system detects that the user tends to associate red objects with positive affective states, the EA would tend to increase the amount of red objects in the virtual environment. An introduction to EA can be found in Eiben and Smith [98].

Additional research indicates that it is possible to create automatically visual compositions in VR using Deep Convolutional Neural Networks (DCNN) [99]. Overall, the process consists of merging features from two images to create a third image. This approach could be combined with EDPCG [96, 97] to create personalized affective visualizations. The process would involve at least three steps: (1) Create a set of VR content that all users will observe and use that content as a baseline. This initial set of content could be developed following the guidelines described in Table 2.1); (2) Capture user feedback about the visual stimuli to establish the affective state that each user associates with each piece of VR content; And (3) use DCNNs to merge features of the initial content onto new, personalized VR content.

2.3 Assessment of affective states

Users' feedback should be captured using methods that do not interrupt the VR experience, such as body movements or electrophysiological signals, similar to Georgiou and Demiris [100]. Methods for assessing affective states can be grouped into three categories: self-report questionnaires, behavioral measures, and electrophysiological signals. Each method has advantages and disadvantages that will be discussed below.

2.3.1 Self-reports

Self-reports allow participants to evaluate their affective state by answering a series of questions. They can be used to verify the accuracy of the acquired information through other methods, such as behavioral and electrophysiological signals. Data collected through self-reports are often used as ground truth in the field of HCI [101, 102].

In general, self-report measures are relatively easy to implement in a computational system because they only require displaying a series of questions. Unlike behavioral and electrophysiological methods, self-reports are considered a direct measure because they allow asking participants directly about their mental states [12]. However, they are susceptible to be biased by rational processes. For example, participants who believe that it is expected from them to respond in a certain way might adjust their responses to fulfill that expectation, causing a phenomenon known as experimenter bias [103]. Some available tools for the assessment of affective responses are the Positive and Negative Affect Schedule (PANAS) [104], Self-Assessment Manikin (SAM) [105], and Pick a Mood (PAM) [106]. The PANAS consists of 20 words related to negative and positive feelings (ten negatives and ten positives). Participants use those words to report their affective state. Each word can receive a rating from one to five.

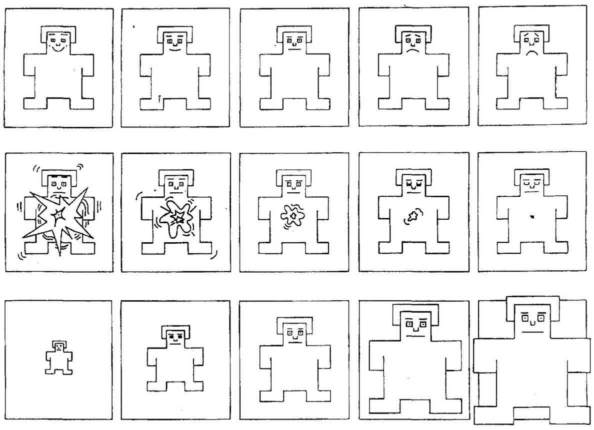


Figure 2.1: From top to bottom: valence, arousal, and dominance scales of the Self-Assessment Manikin (SAM). Taken from Bradley & Lang [105]

The SAM [105] is an instrument that uses three scales: valence (pleasant/unpleasant), arousal (tension/relaxation) and dominance (inhibition/uninhibition). Each scale has five pictograms. Participants can select the blank spaces between each pictogram to indicate intermediate states. Therefore, answers to each scale can take values between 1 and 9 (see Figure 2.1). Given that this instrument is based on dimensions, it is compatible with dimensional models of affect. The SAM [105] is one of the most established instruments for assessing affect (over 7.000 citations) and has been used for the development of batteries of stimuli with emotional content, such as the International Affective Pictures System (IAPS) [69] and the DEAP dataset [23].

On the other hand, the PAM [106] is based on discrete states. Therefore, it is compatible with discrete models of affect. This instrument also uses pictorial cues to assess participant's states. There are eight mood types plus a neutral one: excited, cheerful, relaxed, calm, bored, sad, irritated, and tense. There are three characters for each state: a man, a woman, and a robot (gender-neutral character). In comparison to the SAM [105], PAM's characters [106] are more similar to a real human being (see Figure 2.2), which might be an advantage because it could be easier for participants to feel identified with the characters of the PAM [106]. The PAM [106] has been used to understand how to design objects and experiences that could stimulate mood regulation [42], analyze the effect of immersive virtual environments on gaming Quality of Experience (QoE) [107], and analyze whether the effect of color on affective states varies across different VR rooms [82].

2.3.2 Behavioral measures

Behavioral measures allow inferring affective states from observable conducts, such as body movements [108, 109, 110], voice patterns [111, 112], and facial expressions [63]. During an

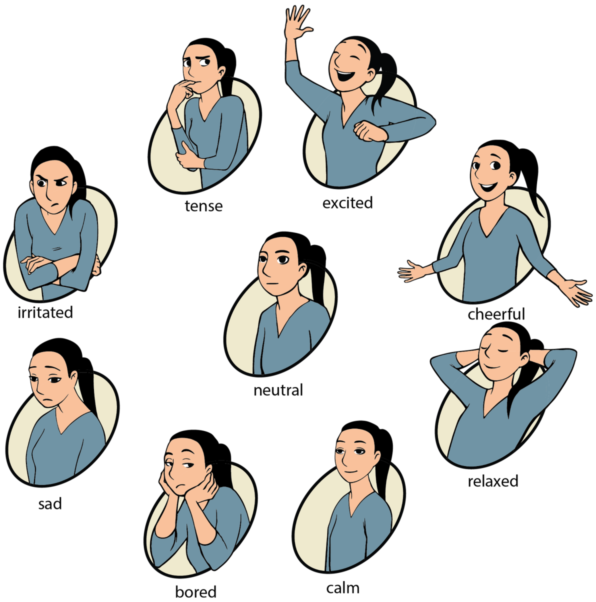


Figure 2.2: Female character of Pick A Mood (PAM), taken from Desmet et al. [106]. Eight discrete states are represented, plus a neutral one.

experiment conducted by Bull [109], participants listened to a series of audio recordings with emotional content while their body movements were videotaped. Results suggested that sadness is associated with dropping the head while boredom is related to leaning the face in one hand. Recent research indicates that it is possible to infer arousal from body movements in virtual reality users [113, 110]. In general, faster body movements are associated with higher arousal.

It is possible to automatically analyze users' affective states based on their voice patterns [114]. Usually, a set of features are defined and used to build a classification model. Some of the features used for automatic speech emotion recognition are pitch, loudness, and tempo [114, 115]. This approach is coherent with evidence suggesting that changes in vocalization patterns have an effect on the affective evaluation of speech [111, 116].

As mentioned in Section 2.1, facial expressions are associated with affective states [63]. These expressions can be analyzed visually and described in terms of the Facial Action Coding System (FACS) [117]. The FACS is an instrument that describes the movements of the facial muscles. Each movement is defined as an Action Unit (AU). Facial expressions can be described as a combination of a subset of all the Action Units defined in the FACS (Ekman and Friesen, 1976). In a study conducted by Porcu et al. (2020), AUs were used for real-time analysis of the facial expressions of video streaming users. Additional studies suggest that human facial expressions can be collected using crowdsourcing techniques [118], and its analysis can be optimized using statistical models that adapt automatically to the characteristics of the data [119]. However, facial recognition with camera sensors might be challenging to implement in VR because the Head-Mounted Display (HMD) covers the user's face. HMDs are the most common device used for interacting with a virtual environment. An alternative that does not require to place a device on users' face are Cave Automatic Virtual Environments (CAVEs). However, CAVEs are not accessible to most users because they are more expensive and more difficult to operate than an HMD. Therefore, facial electromyography (fEMG), a technique introduced in the following section, might be more suitable for capturing VR users' facial expressions [53].

2.3.3 Electrophysiology

This section describes methods to infer affective states in terms of the two dimensions of the CMA [65]: valence and arousal (see Section 2.1.2). There are many approaches for affect detection using electrophysiological signals that are not based on the CMA [65] and are not included in this thesis. Electrophysiological methods allow measuring changes in the electrical potentials of the body. Usually, facial electromyography (fEMG), electrocardiography (ECG), and electroencephalography (EEG) sensors are used to record facial muscle, heart, and brain activity, respectively. In this section, methods that can be used to infer arousal will be presented first, followed by methods that can be used to infer valence.

Arousal can be inferred from features extracted from ECG signals. The features can be extracted in the time and frequency domains. Extracting time-domain features usually requires calculating the time lapse between the peaks (known as the R-peaks) of the ECG signal, yielding the RR-Intervals (RRIs). These RRIs are used to calculate time-domain features, such as the root mean square of successive differences (RMSSD) and the standard deviation of normal intervals (SDNN) [1]. These features are associated with Heart Rate Variability (HRV). It has been found that higher HRV is associated with higher arousal [62, 120]. It is possible to obtain information about cardiovascular activity in the frequency domain as well, calculating the LF/HF ratio. The low-frequency component (LF) (0.04 to 0.15 Hz) is associated with parasympathetic activity, while the high-frequency component (HF) (0.15 to 0.4 Hz) is associated with sympathetic activity [121]. The activation of the parasympathetic system is associated with relaxation, and activation of the sympathetic system is associated with increased arousal. Therefore, increased activity in the HF component indicates higher arousal [122]. Further research has shown that it is possible to infer arousal from EEG signals in VR users employing long short-term memory recurrent neural networks (LSTM-RNN) [123].

A recent study compared the benefits of implementing HRV biofeedback in virtual reality with a traditional HRV biofeedback therapy [36], suggesting that the VR implementation produces more benefits for users in terms of relaxation, self-efficacy, reduced mind wandering, and control of attentional resources. A similar approach was proposed in Blum et al. [18], introducing a breathing biofeedback algorithm. This algorithm combines features extracted from ECG signals with data inferred from diaphragm movements. The experiment was conducted using a chest band (Polar H10), a reliable, relatively inexpensive sensor. Results suggest that the approach proposed by Blum et al. [18] can help to foster more regular and slower breathing in VR users.

Valence can be inferred from EMG and EEG signals. Previous evidence suggests that the Corrugator Supercilii muscle activity (located above the eyebrows) is associated with negative affective states. In contrast, the Zygomaticus Major muscle activity (located in the cheeks) is related to positive affective states [124]. Changes in facial muscle activity can occur without conscious awareness of the participant [60, 125]. However, it might be challenging to implement EMG in a VR system because the Head-Mounted Display (HDM) pressure on the electrodes can create artifacts on the recorded signal.

Asymmetry in the cortical activity of the frontal cortex is also associated with valence. It has been found that positive and negative affective states are processed in the left and right frontal cortex, respectively [126, 61, 127]. Additionally, it has been found that cortical activity decreases as the alpha

[1] Normal intervals consist of RR-intervals that fall within a statistically normal distribution.

power (frequencies between 8 and 13 Hz) increases [128, 129]. Therefore, increased processing of positive stimuli is associated with decreased alpha power in the left frontal cortex (higher activity in the left side of the brain). Similarly, increased processing of negative stimuli is associated with decreased alpha power in the right frontal cortex (higher activity in the right side of the brain) [130, 128, 61]. A similar phenomenon has been observed in the theta band (frequencies between 4 and 12 Hz) [129], indicating that frontal theta asymmetry might be associated with valence as well.

These findings are coherent with results obtained by Reuderink et al. [131] in a study where the brain activity of video game players was recorded using EEG. Participants were asked to report their affective state using the SAM [105] after the game session ended. Results indicated a positive correlation between self-reported valence and alpha asymmetry. Likewise, Koelstra et al. [23] analyzed the brain activity of 32 participants who watched forty musical videos and rated their emotional reactions using the SAM [105]. A positive correlation was found between self-reported valence and alpha power in the right occipital region of the brain.

EEG signals are particularly prone to capture noise (i.e., electrical activity unrelated to mental states). For example, eye movements and eye blinks produce artifacts in the EEG signals and are usually reflected in the activity of the brain's frontal region. In non-stationary VR applications, it is particularly challenging to remove artifacts produced by muscle activity, head movements, or electrical activity from the VR headset [132]. It is possible to remove these artifacts using Independent Component Analysis (ICA). This technique allows identifying the components of an EEG signal that are not produced by brain activity [133]. The maximum number of independent components (ICs) that can be identified using ICA depends on the number of electrodes used. For example, a recording with 32 electrodes will allow the identification of up to 32 ICs. Therefore, increasing the number of electrodes might help identify the artifacts in the signal with more precision. For a complete analysis about using ICA in non-stationary and stationary settings, see Klug and Gramann [132].

An additional challenge is to process the EEG signals in real-time. ICA can be used in real-time [134], but it was not designed for that purpose. An alternative is Artifact Subspace Reconstruction (ASR) [135, 136], a technique designed for online artifact removal. ASR uses data recorded from the participant as a baseline. Then, Principal Component Analysis (PCP) is applied to identify the EEG channels that contain artifacts. The data of the corrupted channels are reconstructed using the baseline data as a reference. There is software available that can facilitate the implementation of ASR, such as BCILAB [137] and Neuropype (Intheon Labs, California).

2.4 Brain-Computer Interfaces

The implementation of electrophysiological signals in VR systems leads to the development of interfaces that allow interpreting users' brain activity as computer commands [138, 139]. One of the basic assumptions underlying the development of Brain-Computer Interfaces (BCIs) is that mental processes originate in the brain. However, there are BCIs that measure electrophysiological responses in other places of the body [140], such as the heart and face, because processes that originate in the brain can produce changes in the activity of other body parts.

There are different techniques for measuring brain activity that can be used for the development of BCIs. For example, electrocorticography (ECoG), Positron Emission Tomography (PET), and functional Magnetic Resonance Imaging (fMRI), among others. However, electroencephalography (EEG) is the method most frequently used in BCIs because (1) it provides high temporal resolution (i.e., relatively large amount of data points recorded per second); (2) it does not create health risks for the user because the electrodes can be easily placed and removed from the scalp; (3) can be portable, which is important for applications where the user is moving; and (4) is less expensive than most of the other methods [20].

According to Zander and Kothe [20], there are three types of BCIs: active, passive, and reactive. Active BCIs require the active participation of the user to generate an action. For example, patients who lack motor control can use mental commands to move a wheelchair [141]. Passive BCIs do not require the conscious involvement of the user. They can be used, for example, to analyze the cognitive load of car drivers automatically [142]. Reactive BCIs use mental activity that occurs as a response to external stimuli. An example is a neurofeedback video game where threatening stimuli are presented, and players have to control their anxiety to obtain game score [143]. The typical workflow in a BCI involves at least four steps [20, 127]:

1. **Preprocessing pipeline:** Filter out the signal's noise and keep only the components that reflect brain activity. This process involves (but is not limited to) filtering frequency bands and removing artifacts caused by eye movements or muscle activity. An introduction to signal processing can be found in Unpingco [144].

2. **Feature extraction:** Isolate the information related to the psychological construct of interest based on previous neuroscience studies (see Section 2.3.3).

3. **Classifier definition:** A classification model is build using prerecorded data. The classifier is tested offline, and an estimate of the accuracy of the classification is calculated. In general, classifiers are trained using data that has been previously labeled by humans. Machine Learning algorithms are used to identify patterns in the data that tend to be associated with each label.

4. **Classification application:** The classification is implemented in the BCI to perform instantaneous analysis of the brain activity. The outputs of the classification are used as computer commands.

2.5 Chapter discussion

This chapter analyses previous studies in psychology, electrophysiology and audiovisual design. This interdisciplinary analysis might contribute to understanding how to develop VR systems for affective visualization. These VR systems would involve developing at least two components: a virtual environment and an affect detection technique. The development of both components requires the understanding of theories related to emotion and affect. Therefore, this chapter analyses previous research related to (1) theories of emotion and affect, (2) audio-visual cues associated with affective states, and (3) methods for assessment of affective states.

Studies discussed in Section 2.2 suggest that specific visual and sound cues can represent users' emotions. However, most of these studies were conducted in experimental settings where the stimuli were carefully controlled. It is unclear whether the same psychological responses would occur if a combination of these cues were used simultaneously. For example, a particular combination of colors associated with positive states may result in an unbalanced visual composition that produces negative affective states. Alternatively, there might be motion patterns that are more prone to produce motion sickness in VR users, triggering negative states. Furthermore, the sense of novelty that a VR system produce in new users might influence the emotions triggered by the virtual environment.

Other studies mentioned in Section 2.2 suggest that leftwards linear motion tends to be associated with negative valence [87, 85]. This finding was obtained during experiments conducted in a western society, where time is represented as a progression to the right [145]. Therefore, it is likely that western users associate leftward motion with negative affective states because that type of motion is culturally associated with regressing in time. However, in other cultures, such as the Hebrew culture, people represent time as a progression to the left [145]. Therefore, it is possible that Hebrew users would associate leftward linear motion with positive valence.

Recent studies have demonstrated that affective states can be elicited by triggering psychogenic shivering (PS) [146, 147], using a device that controls the temperature in the upper back of the participants. Additional research indicates that the ability to be empathetic with others' emotions can be influenced by delivering electrical stimulation in the vagus nerve [148], and by inducing affective states in the observer through videos [149]. It remains an open question how to use those findings to develop Mixed Reality (MR) technologies for empathy enhancement, as proposed by Schoeller, et al. [147].

Most of the existing techniques for inferring affective states from electrophysiological signals are based on a small number of discrete states [52, 53]. However, the amount of distinct affective states that can be detected using this approach is limited. Therefore, it might be convenient to formulate affect detection problems in terms of statistical regression. This approach would allow building a model capable of describing affective states in terms of a continuum containing an infinite amount of distinct affective states. Previous studies suggest that it is possible to infer arousal from EEG signals [123] as a continuous variable. Future studies could investigate whether it is possible to use a similar approach to express valence in terms of a continuous variable.

Finally, it is possible to use a programmatic approach to create VR content in real-time, using procedural content generation (PCG) [96, 97, 33, 21]. PCG allows to create content dynamically that adjusts to user feedback. Electrophysiological signals could be used to capture user feedback without interrupting the VR experience. This approach would allow the creation of personalized virtual environments for affective visualization, similar to Kitson et al. [99] or Bermudez i Badia et al., [33].

3

Visual metaphors of affect

This chapter is based on a preprint paper [150].

Is it possible to observe affective states? Previous experiments suggest that it is possible to create audiovisual representations of the users' mental states [151, 152, 153]. This practice, known as brain-painting, has shown that it is possible to use Brain-Computer Interfaces (BCIs) to analyze users' brain activity in real-time, interpret that information as computer commands, and automatically generate audiovisual metaphors of mental states.

This chapter aims to understand how to design visual metaphors of affective states. This challenge is addressed by developing five Virtual Reality (VR) environments for affective visualization. The design of these virtual environments is based on previous studies suggesting that visual cues, such as colors [77, 76], shapes [85, 86, 87] and motion patterns [71, 73] tend to be associated with particular affective states, regardless of individual differences between users (see section 2.2).

The environments were tested in an exploratory experiment where participants indicated the affective state they associated with each environment. Results suggest that the environments represent the affective states that they were expected to represent. The aim is to use these virtual environments in a neurofeedback VR system (see chapter 5). In this system, players are asked to recall four autobiographical memories while their electroencephalography (EEG) signals are analyzed in real-time. The information extracted from the EEG signals will allow to infer the affective states triggered by those memories. The visual stimuli in the VR system will be adjusted accordingly. Previous evidence indicates that this procedure might increase the ability of users to self-regulate their affective states [44, 154, 155].

The development of this system requires visual cues that represent similar affective categories for any player. Otherwise, the visual cues will not accurately inform participants about the affective states they are evoking with their memories. Emotional experiences are subjective, hence players might associate similar visual stimuli with different affective states. For example, a user who suffered a traumatic experience with fire might associate the shape of a bonfire with negative states, while others might associate the same shape with positive experiences.

At the same time, previous evidence (see section 2.2) suggests that it is possible to design visual compositions that represent similar affective states for different users. On the one hand, psychology studies support a user-independent mechanism for processing affective stimuli. Ekman and Friesen

[63] found evidence suggesting the existence of at least six basic facial expressions of affective states (happiness, sadness, boredom, disgust, fear, and anger) that are culture-independent. Further studies indicate that facial muscles are automatically activated when negative and positive stimuli are observed [124, 60]. Additional evidence suggests that negative and positive affective states are processed in the right and left frontal cortex, respectively [130, 61, 128]. Taken together, these studies suggest the existence of a psychological mechanism related to the perception of affective stimuli that operates similarly in different users.

On the other hand, previous studies in visual perception suggest the existence of visual cues that tend to be associated with particular affective states (see section 2.2.). It has been found that features such as shape, color, and visual complexity tend to be associated with affective responses. Rounded objects tend to be associated with higher valence and lower arousal than sharp objects [73, 72, 71]; rounded lines tend to be associated with higher valence than straight lines [73, 72]; and extremely complex objects tend to be more pleasant (higher valence) than objects with middle-level of visual complexity [74].

Abstract figures are not directly associated with real-world contexts. Hence using abstract figures for affect visualization might help reducing associations with personal experiences, while avoiding biases caused by individual differences between users. On that account, the virtual environment presented in this chapter was developed using abstract figures.

Since VR is more immersive than other technologies [13, 12], it is more likely that users will engage with a VR system, as compared to an experience based on traditional technologies that use 2D displays. Therefore, VR is an ideal technology to create a biofeedback therapy experience for training affective self-regulation.

3.1 Virtual Environment

Five virtual environments were created using Unity. The design process was based on graphics design guidelines found in previous studies [77, 87, 71, 85, 86, 156, 79, 72, 73, 85, 86] and on the criteria of a researcher with training in graphics design (Andres Pinilla). The five environments had a similar structure: a flat landscape with a particles systems surrounding the participant's visual field. The graphical properties of the particles system in each environment were different. Each environment was intended to represent one of the following types of affective states: (1) High Arousal, High Valence; (2) Low Arousal, High Valence; (3) Low Arousal, Low Valence; (4) High Arousal, Low Valence; and (5) Neutral (experimental control). Details about the design of each environment are described below.

3.1.1 Exciting environment: High Valence, High Arousal

A color palette with saturated, bright colors was used for the exciting environment (see Figure 3.1) because previous studies suggest that these colors tend to be associated with high valence (positive affective states) [77]. The particles system was in front of the participant's visual field and moved towards the user's position in a circular fashion. It was expected that this would lead to higher arousal ratings because circular motion patterns seem to be associated with high arousal [87].

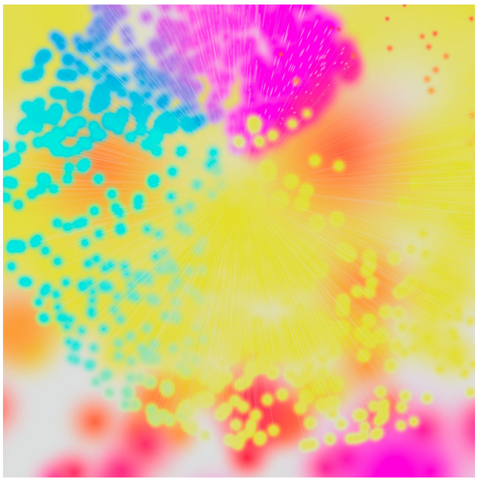


Figure 3.1: Screenshot of the exciting environment.

3.1.2 Calming environment: High Valence, Low Arousal

Previous studies suggest that rounded shapes are associated with high valence and low arousal [71]. Therefore, the calming environment (see Figure 3.2) consists of rounded particles that appeared in the air and shrank gradually until vanishing. Unsaturated blue and purple colors were used to trigger associations with high valence and low arousal [77]. The particles moved slowly upwards, a motion pattern that tends to be associated with low arousal [85, 86]. The participant controlled the position where new particles appeared by moving the head. The landscape resembled a natural field, with green grass on the floor, a blue sky, and stars placed randomly across the space. Given that this scenario resembled a peaceful, natural landscape, it was expected that participants would associate it with positive, relaxing mental states (high valence and low arousal).

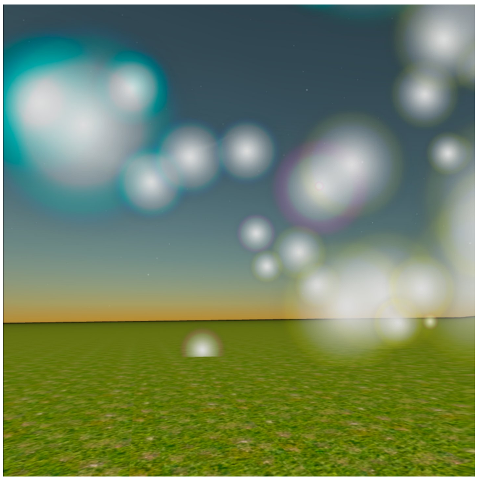


Figure 3.2: Screenshot of the calming environment.

3.1.3 Depressing environment: Low Valence, Low Arousal

The depressing environment (see Figure 3.3) consisted of grey particles. Unsaturated, dark colors were used because previous studies suggest that these are associated with unpleasant affective states [80]. The fog moved following a linear motion because that motion pattern seems to be associated with low arousal [85]. In the background, there was a black sky, which tends to be associated with negative affective states (low valence) [157].

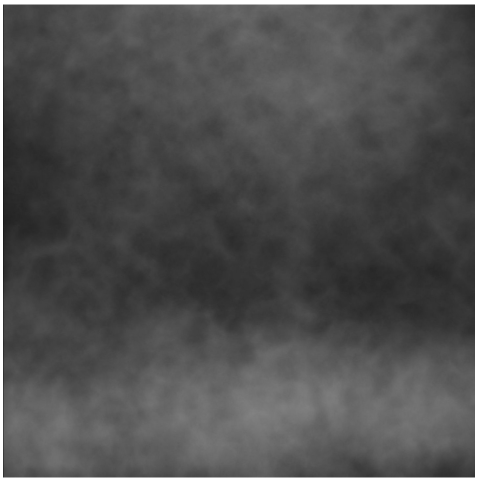


Figure 3.3: Screenshot of the depressing environment.

3.1.4 Stressing environment: Low Valence, High Arousal

Yellow-green colored straight lines were used in the stressing environment (see Figure 3.4) because this color tends to be associated with high arousal and low valence [79, 77]. The lines were moving very fast in multiple directions, crossing paths with each other. The intersection between these lines created angular shapes, which tend to be associated with high arousal [72, 73]. Lines appeared and disappeared very fast, which created the sensation of high speed, which is associated with high arousal as well [85, 86]. The fact that participants had no control over the direction of the lines or the position where they appeared was indented to produce a sense of chaos.

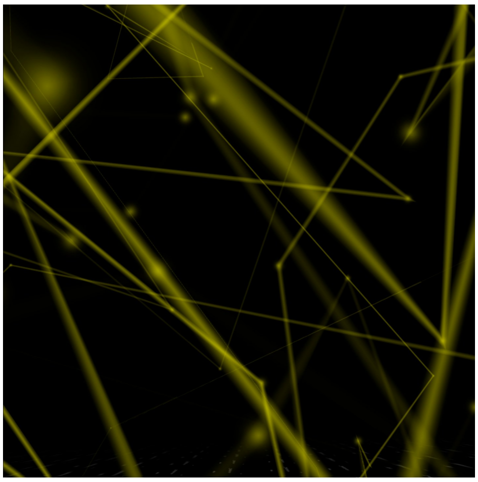


Figure 3.4: Screenshot of the stressing environment.

3.1.5 Neutral environment: Neutral Valence, Neutral Arousal

The neutral environment (see Figure 3.5) consisted of an empty room with white walls. This was the simplest environment but the most difficult to design because it was challenging to find neutral stimuli. An empty white space was used because it represented the concept of *nothingness*, which might be associated with neutrality.

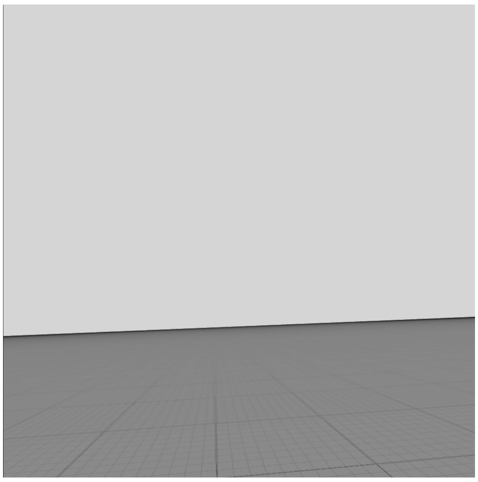


Figure 3.5: Screenshot of the neutral environment.

3.2 Methods

3.2.1 Participants

Twenty-one university students participated in the study (9 women and 12 men). Their age was between 18 and 31 (M = 24.81; SD = 3.57). All participants gave their written informed consent before participating in the experiment. No payment was offered for participation.

3.2.2 Apparatus

The virtual environment was developed in Unity 2020.1.3 and presented in an Oculus Quest, which consists of a HMD (display resolution: 1440 x 1660 pixels) and two controllers (one for each hand). The Self-Assessment Manikin (SAM) [105] was used to evaluate the affective states associated with each virtual environment.

3.2.3 Procedure

After providing informed consent, participants were asked to sit in a chair and the Head–Mounted Display (HDM) was placed in their head. The left and right thumbsticks were used for rotation and walking, respectively. A gaze pointer was used to select buttons inside the virtual environment. The participant moved the head to position the gaze pointer over the desired button in the virtual environment and then pressed the "A" button of the right controller.

The experiment was divided into two phases. In the first phase, participants navigated a practice environment with the same graphical structure as the environments described in section 3.1. The practice environment allowed participants to gain familiarity with the layout of the virtual environment and the controls.

In the second phase, participants explored the five environments presented in Section III. The environments were presented in random order. After each environment, a white screen with the following statement was shown: "Choose the pictogram that better represents the previous environment." Below this statement, one of the rating scales was shown. When the answer was submitted, the following scale was prompted. The scales were prompted one at a time: first valence, second arousal, and third dominance. Labels were placed in the poles of each scale, similar to Lombard et al. [158]. The labels used for the valence scale were "happy" and "unhappy," for the arousal scale were "excited" and "calm," and for the dominance scale were "controlled" and "in control." When the answers to the three scales were submitted, the following environment was shown.

The results of the dominance scale were analyzed but not included in this chapter because the aim was analyzing affective states in terms of the CMA [65], which has only two dimensions: valence and arousal. The responses towards the dominance scale are available in the public repository of the research project [1]. The CMA was used for this experiment because it provides a theoretical framework to analyse affective states that has been used in previous studies in psychology [69, 23].

3.2.4 Ethics

The experiment was approved by the Human Research ethics committee of the University of Technology Sydney, with approval number ETH20-5123.

3.3 Data Analysis

A within-subjects design was used, with *environment* as the dependent variable, with five levels: exciting, calming, depressing, stressing, and neutral. The independent variables were the ratings in the valence and arousal scales.

3.4 Results

A MANOVA was conducted to analyze whether the environment explored by participants produced a significant effect on the ratings in the valence and arousal scales. No outliers were detected. The assumption of normal distribution was violated in 6 of the 10 cells of the experimental design, as assessed by Shapiro-Wilks test ($p < .05$). Despite of this limitation, the analysis was conducted because ANOVA is robust towards violations of normality [159]. The assumption of homogeneity of variances was met for the three dependent variables, as assessed by Levene's test ($p > .05$). There was a statistically significant difference between the evaluations towards each environment, $F(8, 198.000) = 11.1171$, $p < .001$, Wilk's $\Lambda = 0.4762$.

Univariate analyses revealed a statistically significant effect in valence, $F(4, 80) = 17.518$, $p < 0.001$ and arousal, $F(4, 80) = 20.648$, $p < 0.01$. No significant effect was found in the dominance scale, $F(4, 80) = 2.441$, $p = 0.053$. A Tukey's post-hoc test was conducted to analyze the differences in the ratings between each environment on the valence and arousal scales. The valence ratings were significantly higher ($p < .05$) for the exciting ($M = 6.95$; $SD = 2.06$) and calming environment s ($M = 7.81$; $SD = 1.17$), as compared to the ratings of the stressing ($M = 4.9$; $SD = 1.97$) and depressing ($M = 4.59$; $SD = 2.26$) environments (see Figure 3.6). The ratings towards the neutral environment ($M = 5.38$; $SD = 2.11$) were significantly lower ($p < .05$) than the ratings of the calming environment, but not significantly different to the other environments ($p > .05$).

[1] https://osf.io/6wc5e/

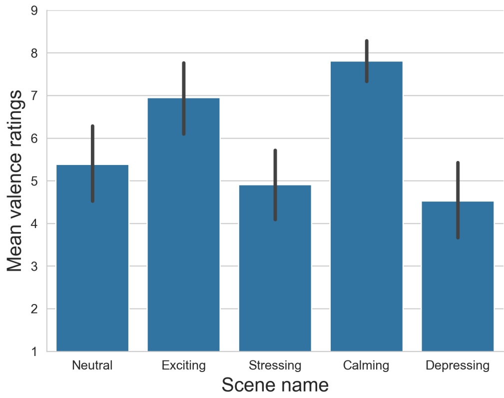


Figure 3.6: Mean valence ratings of each environment in the valence scale of the SAM. Error bars depict CI, 95%.

The arousal ratings (see Figure 3.7) were significantly higher ($p < .05$) for the exciting ($M = 6.14$; $SD = 2.46$) and stressing environments ($M = 6.33$; $SD = 2.33$), as compared to the arousal ratings of the calming ($M = 3.29$; $SD = 2.49$) and depressing ($M = 3.59$; $SD = 1.76$) environments. The arousal ratings towards the neutral environment ($M = 2.9$; $SD = 1.87$) were significantly lower than the ratings of the exciting and stressing environments ($p < .05$), but not significantly different to the calming and depressing environments ($p > .05$).

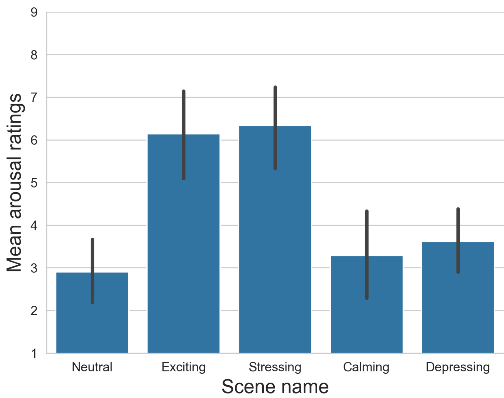


Figure 3.7: Mean arousal ratings of each environment in the valence scale of the SAM. Error bars depict CI, 95%.

In summary, results indicate that the exciting and calming environments are associated with higher valence than the stressing and depressing environments. Similarly, the exciting and stressing environments are associated with higher arousal than the calming and depressing environments. The neutral environment is associated with neutral valence (i.e., higher valence than the stressing and depressing environments, and lower valence than the exciting and calming environments). At the same time, the arousal rating of the neutral environment is similar to the calming and depressing environments, and lower than the arousal in the exciting and stressing environments.

The results obtained with the post-hoc tests indicate that the exciting environment is associated with high valence and high arousal; calming with high valence and low arousal; depressing with low valence and low arousal; stressing with low valence and high arousal; and neutral with neutral valence and low arousal.

The mean ratings indicate something similar. The mean ratings in the valence scale for exciting, calming and neutral environments are above 5 (high valence), and in the stressing and depressing environments are below 5 (low valence). Similarly, the mean ratings in the arousal scale for the exciting and stressing environments are above 5 (high arousal), and in the calming, depressing and neutral environments are below 5 (low arousal). Mean ratings below five are considered "low" and mean ratings above five are considered "high" because the scales range from 1 to 9. These results suggest that the type of affective states associated with each environment (see Table 3.1) are coherent with their graphical properties (see section 2.2).

Table 3.1: Type of valence and arousal associated with each environment.

Environment	Valence type	Arousal type
Exciting	*High*	*High*
Calming	*High*	*Low*
Depressing	*Low*	*Low*
Stressing	*Low*	*High*
Neutral	*Low*	*Low*

3.5 Discussion

This chapter presents five virtual environments: exciting, calming, depressing, stressing, and neutral. These environments were tested in an exploratory experiment with twenty-one participants who evaluated the affective states that those environments represented. Overall, the results suggest that the virtual environments can be used to visualize four types of affective states, covering the four quadrants of the CMA [65]. The objective was to use these environments to provide affective visual feedback in VR, as part of a neurofeedback therapy for training affective self-regulation.

The mean valence ratings of the stressing and depressing environments were higher than expected (slightly below 5, near the middle of the valence scale). Therefore, they do not seem to be associated with extremely negative states. In fact, the valence ratings of the depressing and stressing environments were similar to the valence ratings of the neutral environment. It is possible that the stressing and depressing environments were aesthetically appealing for participants. Since aesthetics has a role in valence ratings [160], it is likely that participants' aesthetic judgments biased the evaluation of these two environments. It remains an open research topic analyzing how to design abstract visual stimuli that can be used to represent extremely negative states. One possible approach would be to develop virtual environments with aesthetically unappealing stimuli.

The arousal ratings in the neutral environment were similar to the calming and depressing environments. However, previous studies have shown that faces associated with neutral affective states tend to obtain low arousal ratings [161]. Therefore, the arousal levels associated with the neutral environment seem to be consistent with the arousal levels typically associated with neutral faces.

The design of all the environments had a similar graphical structure: a flat landscape with a particles system (abstract figures) placed inside the participant's visual field. Using the same visual structure in all the environments helped to control the variables that could affect participants' ratings, allowing to make a valid comparison between participants' ratings towards each environment. At the same time, this design approach limited the level of realism of the virtual environments. There were no mountains, trees, animals, or other objects that could produce biases in the participants' evaluations.

The results are consistent with previous findings indicating that colors [77, 76], shapes [85, 86, 87] and motion patterns [71, 73] tend to be associated with affective responses. Moreover, it has been shown that combining those visual cues into a single graphical system is possible. This graphical system could be used to create a more complex instrument for affective visualization. For example, using Procedural Content Generation (PCG) to automatically create figures whose graphical parameters are associated with particular affective states, similar to Semertzidis et al. [21]. The variables that would be controlled in the virtual environment would be the visual features identified in section 2.2 (see Table 2.1).

The visual compositions presented in this chapter are not entirely detached from real-world contexts, but aim to produce minimal associations with them. It might be difficult to imagine a visual composition detached from real-world images because it is often assumed that the images people create in their minds are inspired by images previously seen in the physical world. However, there is evidence that refutes this assumption. A previous study suggests that congenitally blind people have dreams with visual content [162]. Therefore, mental imagery is a psychological process that can occur independently of visual perception. Consequently, it might be possible to create abstract representations of affective states, entirely detached from visual references to real-world contexts.

The experiment presented in this chapter was intended to analyze whether there was an association between the environments' graphical properties and the affective states they represent. However, it is not clear whether the visual stimuli influenced participants' affective states. Future studies could address this question by asking participants how do they feel before and after navigating each environment.

4 Affect detection technique

This chapter proposes a technique for near real-time affect detection based on electroencephalography (EEG) signals. The proposed technique was tested with two datasets: The DEAP dataset [23] and data collected during a pilot experiment conducted by the author of this thesis. The first dataset describes affective states in terms of the Circumplex Model of Affect (CMA) [65], while the second describes affective states in terms of the Evaluative Space Model (ESM) [66].

The CMA [65] consists of two dimensions: valence and arousal. One of the fundamental assumptions of this model is that an increase in positive activation implies a decrease in negative activation, and vice versa. Hence the two poles of the valence dimension: pleasant and unpleasant. The Evaluative Space Model (ESM) [66] disputes this assumption proposing the existence of three affective dimensions: negativity, positivity and net predisposition. The ESM [66] allows to represent affective responses with simultaneous positive (pleasant) and negative (unpleasant) activation [163] because both constructs are represented in separate dimensions.

Recent studies support the hypothesis that affective responses do not necessarily fit within the structure of a two-dimensional coordinate system [163]. Yet, it is not clear how to infer users' affective states in terms of a coordinate system that does not assume the existence of a bipolar valence (pleasure-displeasure) continuum.

The first section of this chapter aims to analyse affective states in terms of the CMA [65], using data taken from the DEAP dataset [23]. This dataset contains two versions of the same data: preprocessed and raw data. The preprocessed version is used in this chapter because it is assumed that it has a high signal-to-noise ratio. Therefore, the results obtained with the preprocessed data can be used as a benchmark for subsequent analysis.

The second section describes a pilot experiment in VR where participants observed a subset of the videos used by the authors of the DEAP dataset [23]. Participants were asked to rate their affective response towards the videos in terms of the three dimensions of the ESM [66], while their cortical activity was recorded using EEG sensors. The EEG signals were analysed in a offline setting, emulating the process that would be conducted during near real-time analysis.

Previous evidence indicates that it is possible to infer some characteristics of users' affective states by analysing electrocardiography (ECG) [62, 121, 122], electromyography (EMG) [124, 60] and electroencephalography (EEG) signals [123, 61, 126, 164, 130, 128]. There is evidence suggesting that ECG signals are associated with the intensity of the emotional response [62], EMG signals are associated with the type of the experienced emotion (pleasant vs. unpleasant) [124], and EEG signals are associated both with the intensity and type of emotion experienced by the participant [126, 58].

Further studies indicate that asymmetry in the activity at the frontal region of the brain is associated with valence [130, 126, 61], while activity in the parietal region of the brain is associated with arousal [164, 123]. Some studies suggest that valence is associated with frontal alpha asymmetry [126, 61], while other studies indicate that it is associated with frontal theta asymmetry [129]. Similarly, there is evidence indicating that arousal is associated with parietal alpha power [164], while other studies suggest that it is associated with parietal theta power [123]. Yet, there are other power bands, such as the gamma band, that might be involved in emotional responses as well [23]. Therefore, in this chapter EEG signals are analysed at the five frequency bands commonly used in EEG studies: delta (0–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz) and gamma (30–100 Hz) bands.

4.1 Affect detection in terms of a two-dimensional model of affect

4.1.1 Dataset description

The DEAP dataset is a public database that contains subjective (self-reports) and objective data (physiological recordings) from 32 participants who rated their emotional response towards 40 musical videos. The duration of each video was 60 seconds. Participants were asked to evaluate the emotions that they felt while observing each video using the valence and arousal scales of the Self-Assesment Manikin (SAM) [105]. The participants' cortical activity was recorded using EEG sensors placed at 32 locations, according to the 10-20 international system. Data of the electrode sites F3, F4, P3 and P4 are used in this chapter, similar to Huster et al. [61].

According to Koelstra et al. [23], the steps conducted to preprocess the EEG signals of the DEAP dataset[1] were:

- Down-sample to 128Hz for increasing computer processing speed.
- Remove EOG artifacts, using a Blind Source Separation (BSS) technique [165].
- Apply bandpass filter to remove frequencies below 4.0 Hz and above 45 Hz.
- Common-average referencing.
- Extract epochs. Each segment corresponds to each video observed by each participant. There are 32 participants and each participant observed 40 videos. Thus, there are 1280 segments in total.
- Remove the baseline from the 3 seconds before the start of each epoch.

The authors of the DEAP dataset [23] conducted an online study where participants indicated the type of valence and arousal associated with each video. Based on the results of that study, one of the following labels were assigned to each video:

[1] Details about the DEAP dataset are available at: http://www.eecs.qmul.ac.uk/mmv/datasets/deap/readme.html

1. High Arousal, High Valence
2. Low Arousal, High Valence
3. Low Arousal, Low Valence
4. High Arousal, Low Valence

4.1.2 Feature extraction

Feature extraction was conducted by the author of this thesis in Python 3.7.5. The electrode sites chosen for calculating frontal asymmetry were F3 and F4, while the electrodes chosen to estimate parietal power were P3 and P4, similar to previous studies [61, 123]. Features were extracted following these steps:

1. **Power extraction:** Welch's method was used to extract the power at each power band (delta, theta, alpha, beta, gamma) at each sample of each epoch. An 8 seconds Hann window was used for the Welch's Method. This process was conducted at electrode sites F3, F4, P3 and P4.
2. **Calculate relative power:** Divide the power in each power band by the total power, yielding the relative power for each power band.
3. **Extract relative frontal alpha and theta asymmetry:** Subtract the value from each power band at F3 from F4.
4. **Extract relative parietal alpha and theta power:** Calculate the mean value from each power band at electrode sites P3 and P4.
5. Down-sample the extracted features to 1 Hz to increase computer processing speed.

4.1.3 Data analysis

Linear mixed-effects models were used to analyse the relationship between the features extracted from the EEG signals and the participant's evaluation of the videos. The final aim was to select the features that provide the most valuable information about the affective states of the users. The selected features will be used in the following section to build Machine Learning classification models.

The interaction between each feature and each affective dimension was analyzed in a separate linear mixed-effects model. Each feature was extracted at five power bands. Thus, each power band was analysed in a separate model. This was intended to discern the statistical interactions observed in one power band from the statistical interactions observed in the other power bands. Therefore, the analysis was repeated once for each possible combination of feature (frontal asymmetry or parietal power), affective dimension (arousal or valance) and power band (delta, theta, alpha, beta or gamma). Hence twenty linear mixed-effects models were built (2 features x 2 dimensions x 5 power bands).

The dependent variable in all the models was the rating towards the videos. The dependent variable was the negativity, positivity or net predisposition rating. Frontal asymmetry was the fixed effect in half of the models, while parietal power was the fixed effect in the other half. Gender was

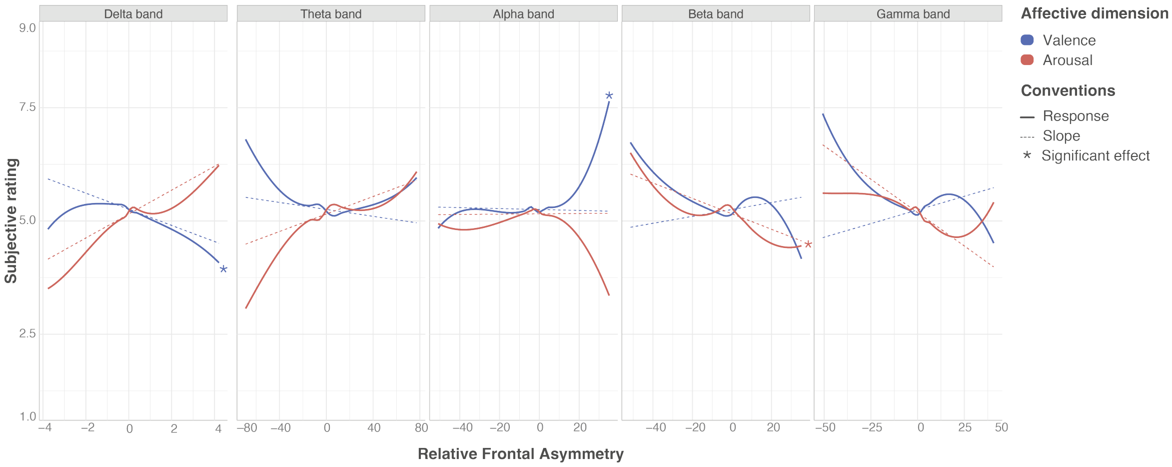


Figure 4.1: Interaction between relative frontal asymmetry (fixed effect) and subjective evaluation of the videos in the valence and arousal dimensions. Each solid line represents a linear mixed-effects model. Dotted lines represent the slope of the corresponding model. This figure does not contain information related to the fixed effects for gender nor the random effects.

the second fixed effect in all the models, because previous studies suggest that women tend to produce stronger responses towards emotional stimuli than men [166].

Three random effects were defined in all the models: video, participant and timestamp. These random effects were chosen based on three assumptions: (1) The electrophysiological responses towards each video are different; (2) Each participant responds differently towards each video; (3) Electrophysiological responses vary across time. Thus, the random effects were the intercepts for video, participants and timestamp, as well as by-video, by-participant and by-timestamp random slopes. The analysis was conducted using R Studio [167] and the lme4 package[168].

Each model was compared against a reduced model that had the same parameters as the full model, but did not contain the feature extracted from the EEG signals. *P*-values were obtained by likelihood ratio tests of the full model against the reduced model. Therefore, a significant *p*-value indicates that the full model (i.e. the model where the EEG data is included) fits the data significantly better than the reduced model.

4.1.4 Results

Visual inspection of Q-Q plots indicated that the assumption of normal distribution was met in all models. Likewise, visual inspection of residual plots indicated that the assumptions of linearity and homoscedasticity were met in all models.

The likelihood ratio tests indicated a statistically significant interaction between valence ratings and frontal asymmetry in the alpha and delta bands, as well as between valence ratings and parietal activity in the gamma band. Additionally, a significant interaction was found between arousal ratings and frontal beta asymmetry. All the other likelihood ratio tests indicated absence of statistically significant interactions (see Table 4.1). Fixed effects are depicted in Figures 4.1 and 4.2.

4.1.5 Classification models

The features used to build the classification models were selected based on the results obtained with the likelihood ratio tests (see Table 4.1). Therefore, frontal alpha asymmetry, frontal delta asymmetry

Table 4.1: Results of the likelihood ratio tests comparing full models against reduced models. Each linear mixed-effects model (full model) was compared against a copy of itself that did not contain the fixed effect of the feature extracted from the EEG signals (reduced model). A significant effect ($p < .05$) indicates that the full model fits the data significantly better than the reduced model.

Feature	Dimension	Power band	*P*
Frontal asymmetry	Valence	Delta	<.001*
		Theta	.227
		Alpha	<.001*
		Beta	1
		Gamma	1
	Arousal	Delta	.518
		Theta	.111
		Alpha	.254
		Beta	<.001*
		Gamma	.283
Parietal power	Valence	Delta	.815
		Theta	1
		Alpha	1
		Beta	1
		Gamma	<.001*
	Arousal	Delta	.304
		Theta	1
		Alpha	1
		Beta	.365
		Gamma	1

* Statistically significant effect.

and parietal gamma power were used to build the valence models. Likewise, frontal beta asymmetry was used to build the arousal models. Two algorithms were used: Linear Discriminant Analysis (LDA) and Logistic Regression (LR), similar to previous studies [169, 170].

The algorithms were trained individually for each participant. With each algorithm, one classification model was trained for valence and another for arousal. Thus, for each participant four models were trained: two for valence and two for arousal. Therefore, in total 128 models were trained (32 participants * 2 affective dimensions * 2 algorithms per affective dimension). In each model, the dataset was split into a training set and a test set. The training set contained 20% of the original set, and the test set contained 80% of the original set. The data was randomly allocated to the training and test set. The valence classification models were trained separately from the arousal models because both dimensions are are conceptually independent [105]. The labels for the videos were taken from the DEAP dataset [23] (see section 3.1.1.). In the valence models, the labels of the videos were "high valence" and "low valence." In the arousal models, the labels were "high arousal" and "low arousal.

Four metrics were calculated in each iteration: accuracy, precision recall and F1-score. The means and standard deviations of those metrics were calculated across participants. The resulting

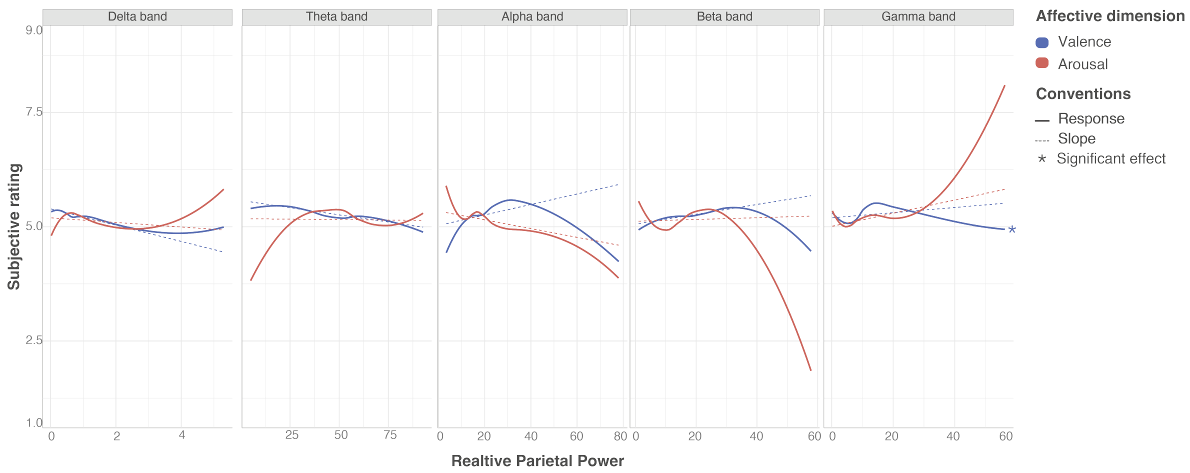


Figure 4.2: Interaction between relative parietal power (fixed effect) and subjective evaluation of the videos in the valence and arousal dimensions. Each solid line represents a linear mixed-effects model. Dotted lines represent the slope of the corresponding model. This figure does not contain information related to the fixed effects for gender nor the random effects.

Table 4.2: Performance metrics of the affect detection models that were built using the preprocessed version of the DEAP dataset [23]. The classification models were trained individually for each participant. The metrics obtained were averaged across participants. Overall, valence models seem to perform better than arousal models. At the same time, LDA seems to perform better than LR .

Dimension	**Algorithm**	**Accuracy**		**Precision**		**Recall**		**F1-score**	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std
Valence	LDA	61.72%	3.87%	62.28%	6.62%	42.78%	19.89%	47.35%	15.13%
	LR	60.68%	4.23%	59.35%	6.04%	38.19%	15.48%	44.69%	14.34%
Arousal	LDA	57.39%	2.66%	58.71%	5.98%	64.88%	24.33%	57.93%	12.04%
	LR	54.54%	3.16%	54.42%	3.22%	55.52%	8.64%	54.7%	5.11%

values are presented in Table 4.2. The metrics obtained for each participant are available in in the public repository of the research project [2].

Overall, LDA models are more accurate than LR models. The mean accuracy of the LDA valence models (M = 61.72%; SD = 3.87%) is 1.04% higher than the mean accuracy of the LR valence models (M = 60.68%; SD = 4.23%). Likewise, the mean accuracy of the LDA arousal models (M = 57.39%; SD = 2.66%) is 6.14% higher than the mean accuracy of the LR arousal models (M = 54.54%; SD = 3.16%). At the same time, valence models tend to be more accurate than arousal models. Interestingly, the standard deviation of the recall metric is higher than the standard deviation of the other metrics, suggesting that recall tends to vary significantly across participants.

4.2 Affect detection in terms of a three-dimensional model of affect

This section contains the results of a pilot study where participants observed a subset of the videos of the DEAP dataset [23] inside a virtual reality home theater. The data was analyzed following the same approach described in section 4.1. Yet, the current section has three main differences with the previous one.

Firstly, in the current section participants' affective states are analyzed in terms of the three dimensions of the Evaluative Space Model (ESM) [66] (negativity, positivity and net predisposition),

[2] https://osf.io/6wc5e/

instead of the two dimensions of the Circumplex Model of Affect (CMA) [65] (valence and arousal). The ESM [66] was used for this pilot experiment because it allows accounting for simultaneous negative and positive activation. Therefore, building an affect detection model based on the ESM [66] might lead to insights about users' affective states that could be obscured by a two-dimensional model [163, 171].

Secondly, the dataset used in the current section is smaller than the DEAP dataset [23]. The analysis presented in the current section was conducted with 15 participants, whose brain activity was recorded at seven electrode sites while observing 16 videos. In contrast, the DEAP dataset [23] was conducted with 32 participants, whose brain activity was recorded at 32 electrode sites while observing 40 videos.

Thirdly, the dataset used in section 4.1 was preprocessed by the authors of the DEAP dataset [23], while the dataset presented in this section was preprocessed by the author of this thesis. The main difference between the preprocessing steps followed by Koelstra et al. [23] and the steps described in the current section lay in the method used for artifact removal. Koelstra et al. [23] used a Blind Source Separation (BSS) method [165] that is not designed for online analysis. In contrast, Artifact Subspace Reconstruction (ASR) [135], a method designed for online artifact removal, is used in the current section.

4.3 Methods

4.3.1 Participants

Twenty-three students from the Technische Universität Berlin participated in the study. One participant was excluded because the file containing the electrophysiological data was corrupted. Two participants were excluded because a flat line was detected in more than 90% of their EEG recordings. Mahalanobis distance revealed outliers in five participants whose data was excluded from further analysis. Therefore, the analysis was conducted with the data of fifteen participants. Their age was between 19 and 58 years old (M = 29.93; SD = 11.49). Six were women, and nine were men. All participants provided written informed consent before participating in the experiment. They received 10€ as compensation.

4.3.2 Virtual environment

The experiment was programmed using the software Psychopy 3.0 [172]. The computer screen was streamed into a Head-Mounted Display (HMD) using the software Virtual Desktop. The virtual environment was a home theatre (see Figure 4.3). Participants remained seated in a chair during the experiment and watched the stimuli on the virtual home-theater screen.

4.3.3 Stimuli

The stimuli consisted of music videos taken from the DEAP dataset [23]. A subset of sixteen videos selected during a previous study [173] were used. Two additional videos from the DEAP dataset [23] were used for training trials. Thus, a total of 18 videos were used. The duration of each video was 60 seconds.

Figure 4.3: Screenshot of the virtual environment used during the experiment. Participants were inside a virtual home theater. The videos were projected on the screen of this home theater using the software Virtual Desktop.

4.3.4 Apparatus

The virtual environment was shown using an HTC Vive. Brain Products amplifier was used for the ECG signals and a g.Tec amplifier was used for the EMG and EEG signals. Recordings from both systems were synchronized using Lab Streaming Layer (LSL). An ECG electrode was placed in each wrist and another in the left ankle. EMG electrodes were placed in the Zygomaticus Major (ZM) and Corrugator Supercilii (CS) muscles, similar to Dimberg et al. [60]. EEG electrodes were placed at F3, F4, P3, P4, T7, T8 and Cz, according to the 10-20 system. The EEG electrode locations were selected based on a previous study conducted by Huster et al. [61] (see Figure 4.4). The reference electrode for EEG was placed in the left mastoid. The ground electrode for EMG and EEG was placed in the right mastoid. The sampling frequency was 5000 Hz for ECG, and 256 Hz for EMG and EEG. Impedance for the EEG signals was below 10 kΩ.

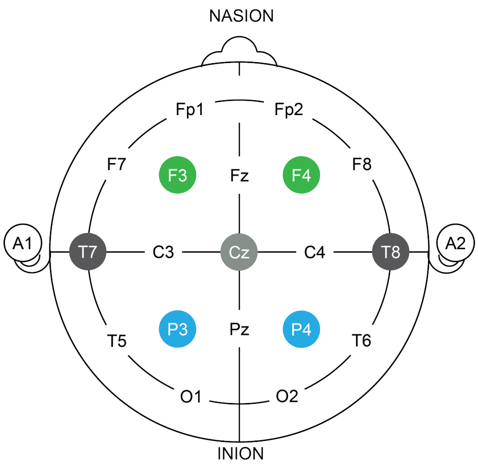


Figure 4.4: Diagram of electrode montage during the pilot experiment. Seven electrodes were placed, according to the 10-20 international system. Electrode sites were F3, F4, P3, P4, T7, T8 and Cz, similar to Huster et al. [61]. Reference and ground electrodes where placed in the left and right mastoids, respectively.

4.3.5 Procedure

Participants signed informed consent and completed a demographics questionnaire. Then, electrodes, HMD and headphones were placed. Electrophysiological signals were visually inspected.

Before starting the experimental task, two practice trials were presented to help participants gain familiarity with the virtual environment and the rating system. Then a 2-minute grey screen with a cross in the center was shown, similar to Koelstra et al. [23]. Finally, 16 trials were presented. Each trial consisted of a 5-second grey screen with a cross in the center, followed by a video. The order of presentation of the videos was randomized. After each video, five questions were prompt. All questions were answered using a slider. Participants used an optical mouse to select their answers in the sliders. The slider for all questions ranged from 1 to 10. In each question, one word was shown in each extreme of the slider, similar to Lombard et al. [158].

The first three questions corresponded to the three dimensions of the Evaluative Space Model (ESM) [66]: negativity, positivity, and net predisposition. The questions were taken from a previous study [149]. The statement of the first three questions was, "how did this video make you feel." In the first question (negativity), the words at the sides of the slider were "1 – Not bad at all" and "10 – Very bad"; In the second (positivity), the words were "1 – not good at all" and "10 – very good"; In the third (net predisposition) the words were "1 – Very relaxed" and "10 – Very restless".

Two additional questions were used to assess liking and familiarity. These questions were taken from Koelstra et al. [23]. The statements of the fourth and fifth questions were "how much do you like this video?" and "how well do you know the video?", respectively. In the fourth question (familiarity), the words placed at the sides of the slider were "1 - Never saw it before the experiment" and "10 - Knew it very well". In the fifth question (liking), the words were "1 - Not at all" and "10 – Very much". Results obtained with the liking and familiarity were analysed but not included in this chapter. The data is available in the public repository of the research project [3].

4.3.6 Signal processing

4.3.6.1 Preprocessing

Anomalies were found in the ECG and EMG data. Therefore, these signals were excluded from further analysis. The EMG data recorded at the Corrugator Supercilii muscle contained noise in all participants, possibly caused by the pressure of the HMD on the electrodes. In the ECG data, the LF/HF ratio [62] was zero for all participants. The ECG and EMG data are available in the public repository of the research project [4].

The following preprocessing steps were conducted in Matlab using EEGLAB 2021.1 [174]:

- **Notch filter:** Remove powerline noise using a notch filter at 50 Hz.
- **Remove bad channels:** Remove channels where a flatline longer than 5 seconds is detected, or whose correlation with nearby channels is lower than 80%.
- **Remove artifacts:** Remove artifacts caused by eye movements, eye blinks and other noise sources using Artifact Subspace Reconstruction (ASR) [135].
- **Re-referencing:** Perform common-average referencing.
- **Band-pass filter:** Apply band-pass filter to remove frequencies below 4 Hz and above 45 Hz.

[3] https://osf.io/6wc5e/
[4] https://osf.io/6wc5e/

- **Extract epochs:** Each video is equivalent to one epoch. Thus, the length of each epoch is 60 seconds.
- **Baseline removal:** Remove baseline of the 3 seconds prior the beginning of each epoch.
- **Down-sampling:** Down-sample to 128 Hz to increase processing speed.

4.3.6.2 Feature extraction

The feature extraction was conducted in Python 3.7.5 following these steps:

- **Power extraction:** Welch's method was used to extract the power at each power band (delta, theta, alpha, beta, gamma), at each sample of each epoch. An 8 seconds Hann window was used for the Welch's Method. This process was conducted at electrode sites F3, F4, P3 and P4.
- **Calculate power ratio:** Divide the power in each power band by the total power, yielding the power ratio at each electrode site during each sample of each epoch.
- **Extract relative frontal alpha and theta asymmetry:** Subtract the value from each power band at F3 from F4. This operation was conducted at each sample of each epoch.
- **Extract relative parietal alpha and theta power:** Calculate the mean value from each power band at electrode sites P3 and P4. This operation was conducted at each sample of each epoch as well.
- Down-sample the extracted features to 1 Hz to increase computer processing speed.

4.3.6.3 Video labels

Three labels were assigned to each video based on participants' evaluations. Each label corresponds to one of the three dimensions of the ESM [66]: negativity, positivity, and net predisposition. For each of these three labels, two categories were defined: high and low. Given that the maximum value of the rating scales was ten, evaluations below five were labeled as "low," and evaluations above five were labeled as "high." Consequently, the labels for the negativity dimension were "high negativity" or "low negativity"; for the positivity dimension, "high positivity" or "low positivity"; and for the net predisposition dimension "high net predisposition" or "low net predisposition."

4.3.7 Ethics

Ethical review and approval were not required for the study following local legislation and institutional requirements. The participants provided their written informed consent to participate in the study.

4.4 Data analysis

Linear mixed-effects models were used to analyse the relationship between the features extracted from the EEG signals and participant's evaluation of the videos. The final aim was to filter out those features that do not provide valuable information about the affective states of the participants. The most relevant features will be used afterward to build classification models.

The analysis was repeated once for each possible combination of feature (frontal asymmetry or parietal power), affective dimension (negativity, positivity and net predisposition) and power band (delta, theta, alpha, beta or gamma). Hence thirty linear mixed effects models were built (2 features x 3 dimensions x 5 power bands).

Three random effects were defined in all the models: video, participant and timestamp. These random effects were defined based on three assumptions: (1) Participant's electrophysiological responses towards each video are different; (2) each participant responds differently towards each video; and (3) electrophysiological responses vary across time.

The analysis was conducted using R Studio [167] and the lme4 package[168]. The dependent variable in all the models was the rating towards the videos in the dimension of interest (negativity, positivity or net predisposition). Frontal asymmetry was the fixed effect in half of the models, while parietal power was the fixed effect in the other half. Gender was the second fixed effect in all the models [166]. The random effects were the intercepts for video, participant number and timestamp, as well as by-video, by-participant and by-timestamp random slopes for the feature of interest (frontal asymmetry or mean parietal power). Each model was compared against a reduced model that had the same parameters as the full model, but did not contain the features extracted from the EEG signals (i.e., the fixed effect). *P*-values were obtained by likelihood ratio tests of the full model against the reduced model.

4.5 Results

Visual inspection of Q-Q plots indicated that the assumption of normal distribution was met in all models. Likewise, the assumptions of linearity and homoscedasticity were met in all models, as assessed by visual inspection of residual plots.

The likelihood ratio tests indicated a significant interaction between the negativity ratings and parietal activity at the theta and beta bands. A significant interaction was found between the positivity ratings and parietal activity in the theta and alpha bands. Additionally, a significant interaction was found between the net predisposition ratings and gamma parietal power. All the other likelihood ratio tests indicated absence of significant interactions (see Table 4.3). Fixed effects are depicted in Figures 4.5 and 4.6.

4.5.1 Classification models

Three models were built for each of the three dimensions of the ESM [66]. Each model was built with one of the following algorithms: Linear Discriminant Analysis (LDA) and Logistic Regression (LR) [169, 175, 170]. The models were built individually for each participant because the aim was to built user-dependent classification models.

The data of each participant was split into a training set and a test set. The training set contained 20% of the original set, and the test set contained 80% of the original set. The data was randomly allocated to each set. Four metrics were calculated for each model: accuracy, precision, recall, and F-1 score. The value of the metrics were averaged across participants. The resulting means and standard deviations are shown in Table 4.4. The metrics obtained for each participant are available in in the public repository of the research project [5].

[5] https://osf.io/6wc5e/

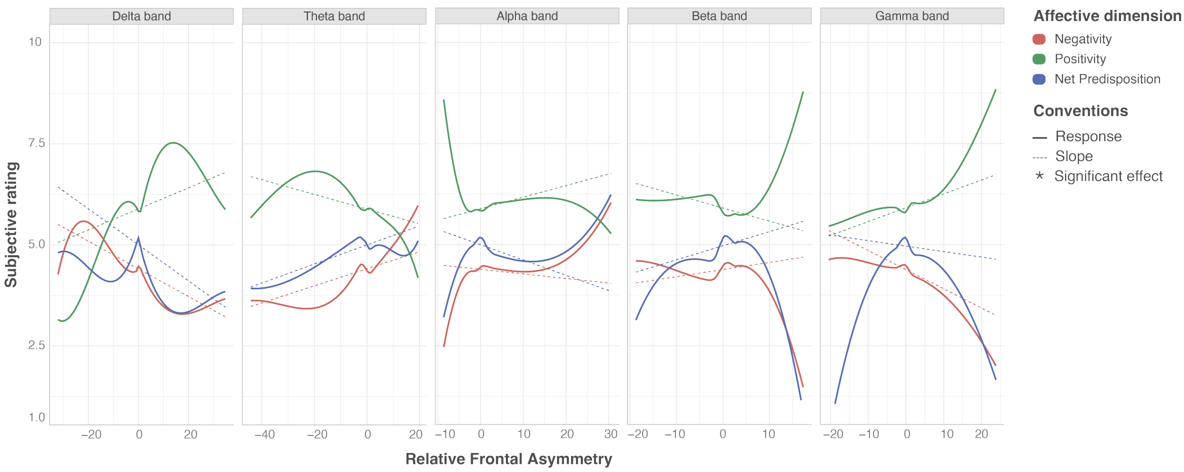


Figure 4.5: Interaction between relative frontal asymmetry (fixed effect) and subjective evaluation of the videos in the negativity, positivity and net predisposition dimensions of the ESM [66]. Each solid line represents a linear mixed-effects model. Dotted lines represent the slope of the corresponding model. This figure does not contain information related to the fixed effects for gender nor the random effects.

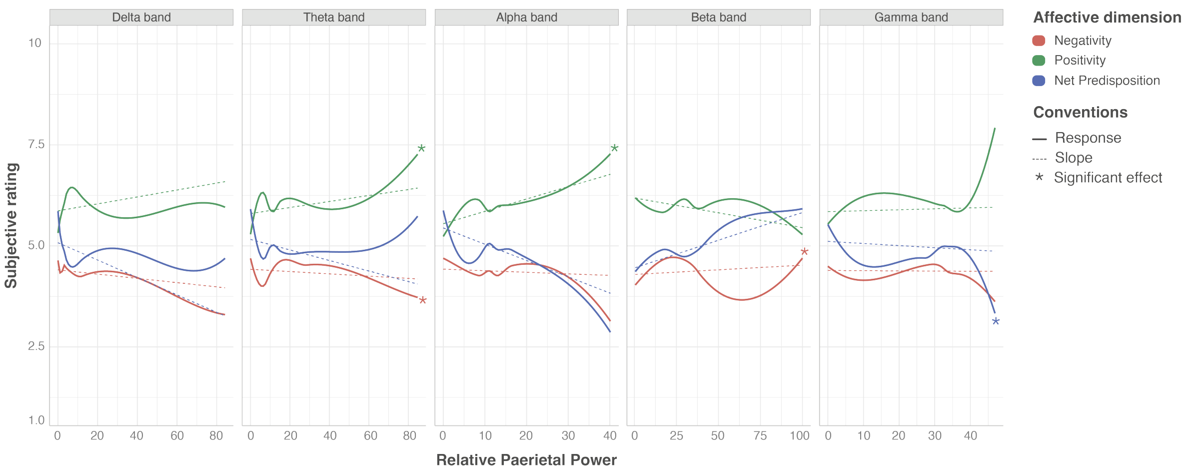


Figure 4.6: Interaction between relative parietal power (fixed effect) and subjective evaluation of the videos in the negativity, positivity and net predisposition dimensions of the ESM [66]. Each solid line represents a linear mixed-effects model. Dotted lines represent the slope of the corresponding model. This figure does not contain information related to the fixed effects for gender nor the random effects.

The features used to built the classification models were chosen based on the results of likelihood ratio tests presented in section 4.4.3 (see Table 4.3. Consequently, the features used for the negativity models were parietal power at the theta and beta bands. The features used for the positivity classification models were parietal activity in the theta and alpha bands. And the net predisposition models were trained with a single feature: parietal activity at the gamma band.

The mean accuracy of the LDA models for negativity ($M = 67.59$; $SD = 7.85$), positivity ($M = 70.05$; $SD = 10.24$) and net predisposition ($M = 64.03$; $SD = 7.25$) was higher than the accuracy of the LR models for negativity ($M = 63.61$; $SD = 9.94$), positivity ($M = 69.1$; $SD = 10.12$) and net predisposition ($M = 61.14$; $SD = 7.38$). In general, precision and recall tend to be lower than accuracy. At the same time, the standard deviation of precision and recall tends to be higher than the standard deviation of accuracy. Therefore, accuracy seems to be a more stable metric across participants than precision and recall. Given that F1-score is a metric derived from precision and recall, the same tendency is observed in means and standard deviations of F1-score.

4.6 Chapter discussion

This chapter analyses how to develop a technique for online affect detection in VR systems. Data from two datasets were analyzed: the DEAP dataset [23] and data collected by the author of this thesis during a pilot experiment in VR. In the first dataset, affective responses are analyzed in terms of the two dimensions of the CMA [65]: valence and arousal. In the second dataset, affective responses are analyzed in terms of the three dimensions of the ESM [66]: negativity, positivity and net predisposition.

In the first section, data was preprocessed by the authors of the DEAP dataset [23]. In the second section, data was preprocessed by the author of this thesis. Other than that, the analysis of both datasets follows the same structure: (1) extract features (frontal asymmetry and parietal power); (2) build a linear mixed-effects model to analyse each extracted feature; (3) select features based on the results of the linear mixed-effects models; and (4) use the selected features to build classification models.

In general, there are two types of BCI classification models: user-dependent and user-independent models. The former are trained with data recorded from the user, while the latter are trained with prerecorded data from multiple users. The affect recognition models presented in this chapter consist of user-dependent models. Thus, the models are tailored to the individual characteristics of each user [176].

The results presented in the first section are consistent with previous studies indicating that valence is associated with frontal alpha asymmetry [61, 126]. At the same time, the results indicate that valence might be associated with frontal delta asymmetry and gamma parietal power as well.

The authors of the DEAP dataset [23] applied a bandpass filter to attenuate frequencies below 4Hz (see section 4.1.1). Therefore, the frequency components of the delta band (0-4Hz) were attenuated. Yet, a significant effect was found between valence ratings and frontal delta asymmetry (see Table 4.1). Therefore, it is possible that the significant effects observed in the delta band are spurious. At the same time, it is possible that the frequency components that are located within the transition band of the filter (i.e., near the 4Hz threshold) are associated with valence. A detailed analysis of the design of the filter used by the authors of the DEAP dataset [23] is required to assess this hypothesis.

A similar finding is observed in the gamma band. A filter was applied to attenuate frequencies above 45Hz in the two datasets analysed in this chapter (see section 4.1.1 and 4.3.6.1). In both datasets, a significant effect was found in the gamma band (30–100 Hz). In the first dataset, the significant interaction was found between valence ratings and parietal gamma power. In the second dataset, the significant interaction was found between net predisposition ratings and parietal gamma activity. Given that the filter was applied at 45Hz, and the gamma band consists of frequency components between 30 and 100 Hz, it is assumed that the statistical effects observed in the gamma band are associated with frequency components that are located between 30Hz and 45Hz.

The dataset obtained during the pilot study is smaller than the DEAP dataset [23]. The total number of trials is around five times smaller in the pilot experiment (15 participants x 16 videos = 240) than in the DEAP dataset [23] (32 participants x 40 videos = 1280). Despite of this limitation, statistically significant interactions were found in the pilot experiment (see Table 4.3). Those findings suggest that (1) it is possible to implement the proposed affect detection technique in a VR system,

(2) it is possible to infer affective states using EEG signals in terms of the three dimensions of the ESM [66], and (3) it is possible to assess affective responses from VR users using only 7 EEG electrodes.

It is expected that the performance of the classification models would increase if they are trained with a larger dataset. There are at least three ways to increase the amount of data collected during future experiments: (1) increase the amount of EEG electrodes, which might help to increase the performance of the online artifact removal algorithm [135], improving the signal-to-noise ratio of the preprocessed data; (2) increase the number of trials and (3) increase the number of participants.

The results obtained with the subjective measures in terms of the CMA [65] and ESM [66] are compared in Appendix A. Overall, this comparison indicates that the ratings reported in the DEAP dataset [23] are similar to the ratings obtained during the pilot study. Thus, the affective states evoked by the videos in the participants of the DEAP dataset [23] were similar to the affective states evoked by the same videos on the participants of the pilot experiment.

The pilot experiment was carried out in a virtual reality environment where participants were not moving. This was intended to reduce the presence of artifacts caused by body movements. Additional research is required to analyze the performance of the proposed technique in non-stationary settings [177]. This could be achieved using a spatial navigation task that involves emotional stimuli, similar to Palmiero et al. [178]. The analysis of the mobile EEG data could be conducted using existing toolboxes for this purpose, such as MoBILAB [179] and BeMoBil [180]. However, those toolboxes are not suitable for online analysis. Therefore, additional research is required to understand how to remove artifacts from EEG signals in an online fashion during non-stationary experiments.

This chapter was focused on EEG signals. However, it is important to consider that EMG signals might have two advantages to identify valence in VR users, as compared to EEG signals. Firstly, EMG signals have been extensively used for studying emotional responses in psychology studies [125, 181, 182], providing robust evidence about their reliability for estimating affective states. Secondly, EMG signals usually require fewer preprocessing steps than EEG signals, which might lead to a simpler, more computationally efficient technique. Future studies could analyze how to build a technique that integrates multiple electrophysiological signals, including ECG and EMG, similar to Cassani et al. [140].

Future studies could analyze how to improve the proposed affect detection technique using Common Spatial Patterns (CSP), a spatial filtering method for maximizing the discriminability between two or more labels in BCI classification problems [183]. A previous study attempted to use CSP to build a valence classifier based on frontal EEG asymmetry [184]. The mean accuracy of the classification models across participants was 46.9%. The mean accuracy of the LDA valence classification model proposed in the current chapter was 61.72 %. It remains an open question whether integrating the affect detection technique proposed in this chapter with the CSP classification technique proposed by Winkler et al. [184] would lead to higher accuracy. Additional studies could analyze the performance of the proposed technique with other datasets, analyzing whether the results reported here are generalizable to a larger population.

Table 4.3: Results of the likelihood ratio tests comparing full models against reduced models. A significant effect ($p < .05$) indicates that the full model fits the data significantly better than the reduced model. The full model contains all the fixed effects and random effects, while the reduced model contains the same parameters, except the fixed effect of the feature of interest.

Feature	Dimension	Power band	*P*
Frontal asymmetry	Negativity	Delta	.289
		Theta	.811
		Alpha	.110
		Beta	.463
		Gamma	.293
	Positivity	Delta	.195
		Theta	.771
		Alpha	.879
		Beta	.285
		Gamma	.523
	Net Predisposition	Delta	.712
		Theta	.401
		Alpha	.426
		Beta	.318
		Gamma	.773
Parietal power	Negativity	Delta	.785
		Theta	<.001*
		Alpha	.254
		Beta	<.001*
		Gamma	1
	Positivity	Delta	.952
		Theta	<.001*
		Alpha	<.001*
		Beta	1
		Gamma	1
	Net Predisposition	Delta	.524
		Theta	.658
		Alpha	.3
		Beta	1
		Gamma	<.001*

* Statistically significant effect.

Table 4.4: Metrics of the classification models that were built using the data collected during the pilot experiment. The models were trained individually for each participant. The mean accuracy, precision, recall and f1-score metrics were calculated across participants.

Dimension	**Algorithm**	**Accuracy**		**Precision**		**Recall**		**F1-score**	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std
Negativity	LDA	67.59%	7.85%	53.27%	35.47%	23.36%	29.07%	34.82%	27.69%
	LR	63.61%	9.94%	32.47%	34.06%	22%	34.48%	33.29%	26.94%
Positivity	LDA	70.05%	10.24%	66.05%	21.32%	90.05%	25.54%	81.37%	8.42%
	LR	69.1%	10.12%	67.47%	14.98%	82.57%	34.79%	71.29%	27.36%
Net Predisposition	LDA	64.03%	7.25%	58.58%	28.38%	55.5%	42.62%	55.18%	28.39%
	LR	61.14%	7.38%	41.39%	31.63%	45.97%	48.54%	56.06%	31.26%

5

Sentui: A reference implementation of a neurofeedback VR system

An experiment with human participants was planned to test the system presented in this chapter. The experiment had to be canceled due to COVID-19 restrictions.

This chapter describes the development of Sentui, a reference implementation of a neurofeedback VR system. The development process was similar to Garcia et al. [16]. Sentui is offered to developers and researchers for use, updates and further development.

The primary purpose of the system is to train users' ability to regulate their affective states voluntarily. Patients who suffer from depression or anxiety might benefit from the system, as it might help to evoke positive emotions. Similarly, people who are at risk of suffering psychiatric disorders might benefit as well because affective self-regulation can help to reduce the risk of suffering mental disorders [185].

The development of the system is based on previous research indicating that it is possible to train affective self-regulation, following a procedure that involves three steps: (1) the user recalls an autobiographical memory that triggers a positive affective state; (2) the system estimates the user's affective state based on his/her brain activity; and (3) visual feedback is provided to the user inside the virtual environment according to the affective state detected [44].

The main characteristics of Sentui are:

- Developed using the Oculus Quest navigation system.
- Representation of affective states in terms of the four quadrants of the CMA [65], using VR stimuli validated in a previous experiment (see chapter 3).
- Affect detection pipeline tested in an offline setting, emulating the steps conducted for near real-time analysis (see chapter 4).
- Two affect detection models: one for valence and another for arousal.

- Data is transferred through the WiFi network using LabStreamingLayer (LSL) [186].
- The system is based on user-dependent classification models (the models are trained with user data).
- The affective state is updated in Unity every second.
- Self-reported affective states are captured with a VR implementation of the Self-Assessment Manikin (SAM) [105] for offline analysis.

5.1 System requirements

Three minimal requirements for the development of the system were identified based on the literature review presented in chapter 2. These requirements are related to (1) the characteristics of the VR stimuli, (2) the characteristics of the affect detection technique, and (3) the phases of the VR experience. Therefore, the following criteria are proposed:

- **VR stimuli:** The VR stimuli should provide visual representations (metaphors) of affective states. The interpretation of these visual representations should be stable across users. This requirement can be achieved by avoiding associations with objects from the physical world, which might help reduce biases caused by the personal experiences of each user. Therefore, four virtual environments were developed using abstract figures. The visual features of the virtual environments are based on the design guidelines summarised in Table 2.1. The virtual environments were validated in the experiment presented in chapter 3. Overall, the results of that experiment indicate that participants tend to associate each of the four virtual environments with one of the four quadrants of the CMA [65].
- **Affect detection:** The affect detection technique should allow inferring the type of valence and arousal experienced by the user in an online fashion and without interrupting the VR experience. This could be achieved using electrophysiological signals [164]. It is possible to use electrocardiography (ECG) signals to infer arousal, electromyography (EMG) signals to infer valence, and electroencephalography (EEG) signals to infer both valence and arousal. For the development of Sentui, EEG signals are used.

  An affect detection technique based on EEG signals should consist of at least three components: (1) a preprocessing pipeline to reduce the presence of spectral components that do not represent brain activity; (2) a feature extraction process to extract the information from the signals that are associated with affective states; (3) a Machine Learning model to automatically assess users' affective states in terms of the two dimensions of the CMA [65]: valence and arousal.
- **Phases of the VR experience:** The VR experience should consist of at least three phases: (1) conduct training trials to collect data of the user and train the affect detection models; (2) induce a baseline affective state in the user; and (3) conduct the neurofeedback task. These phases are based on the experimental design proposed by Li et al. [44].

5.2 Design and development of the system

The system is a closed affect loop consisting of two main modules: the VR content module and the signal processing module. The former was developed in Unity, using the VR stimuli validated in chapter 3. The latter was implemented using NeuroPype (Intheon Labs, California), based on the results presented in chapter 4. The connection between both modules is established through the WiFi network using LabStreamingLayer (LSL). Figure 6.1 describes the architecture of the system.

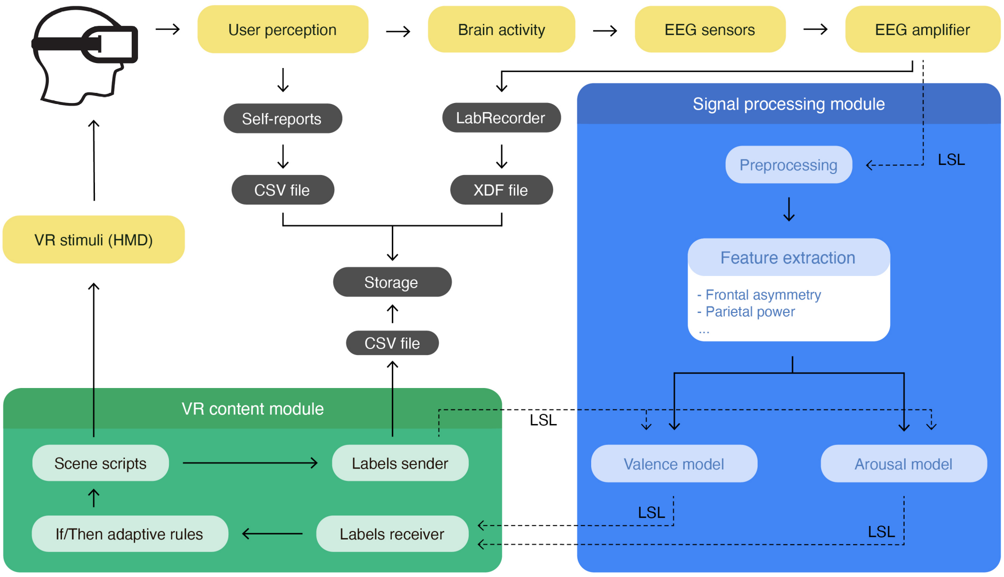


Figure 5.1: Architecture of the proposed reference implementation. The system consists of two modules: the VR content module and the signal processing module. The connection between both module is established using the LabStreamingLayer (LSL) protocol [186].

5.3 The VR content module

Participants are asked to write four autobiographical memories in a paper before using the system, similar to Li et al. [44]. Each memory should be associated with one of the four quadrants of the CMA [65]:

- High Valence, High Arousal: **Exciting memory**
- Low Valence, High Arousal: **Relaxing memory**
- Low Valence, Low Arousal: **Depressing memory**
- High Valence, Low Arousal: **Stressing memory**

There are four virtual environments. Each virtual environment is associated with one of the quadrants of the CMA [65]. A text inside the virtual environment asks the user to recall one of the four memories (see Figure 5.2).

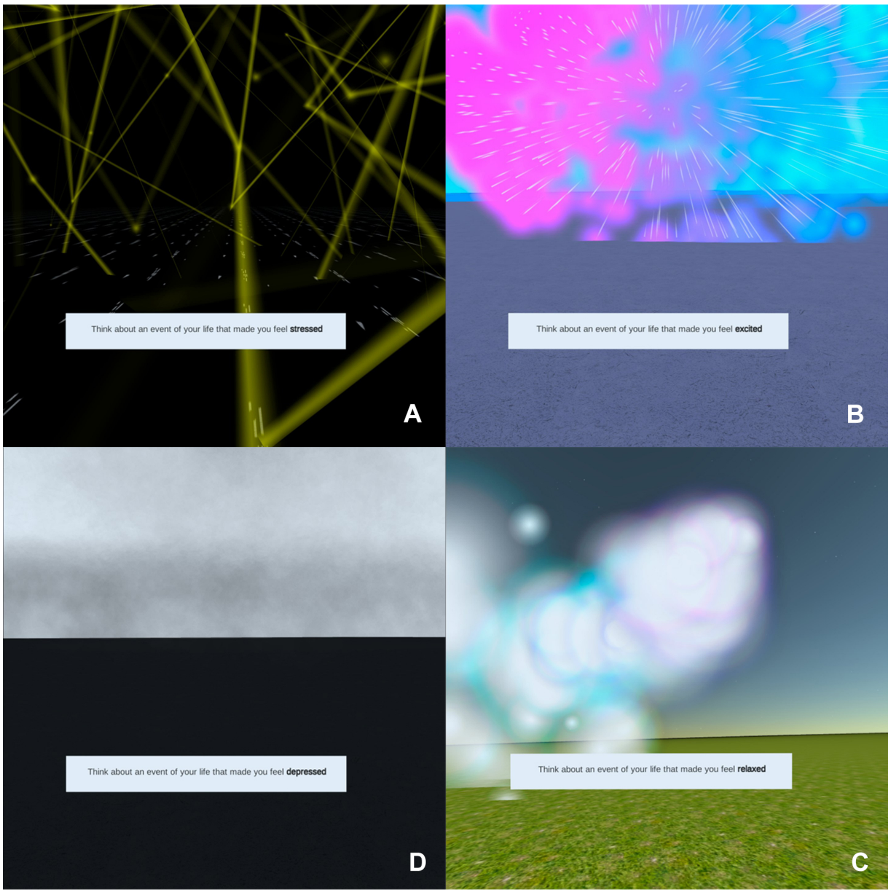


Figure 5.2: Screenshots of the VR scenes used for the VR content module. Each scene is associated to one of the quadrants of the CMA [65]. The association between the scenes and affective states was analysed in chapter 3.

The VR experience consists of three phases, similar to the experimental design proposed by Li et al. [44] (see Figure 5.3). During the first phase, the user navigates five times in each virtual environment for 60 seconds each time. EEG data is labeled based on the type of valence (high valence or low valence) and type of arousal (high arousal or low arousal) associated with each virtual environment. The labels are sent to the signal processing module and used to train the classification models. During the second phase, a baseline affective state is induced. The induction tasks consist of navigating a randomly selected virtual environment for 60 seconds.

During the third phase, three neurofeedback trials are conducted. During the three trials, the user is asked to recall the memory associated with the exciting memory. The virtual environment changes every second according to the affective state detected with the EEG signals. The detected affective state is used to execute if/then commands once per second. If high valence and high arousal are detected, the user obtains game score and the exciting environment is displayed. Otherwise, game score is deducted, and the virtual environment corresponding to the user's affective state is displayed.

Before and after each trial, participants are asked to report their affective state using the SAM [105]. The statement of the question prompt after each memory recall trial is: "Please choose the

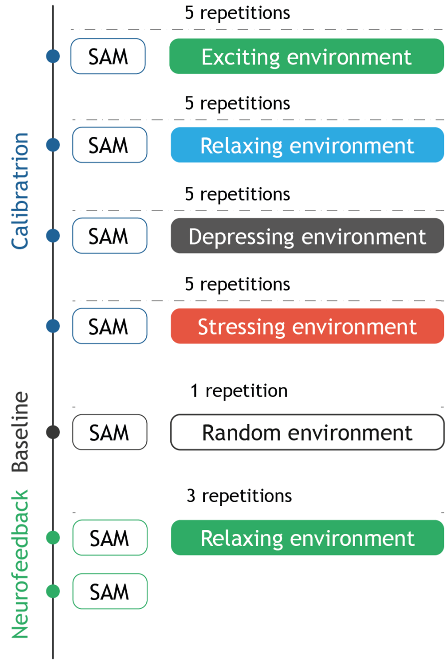


Figure 5.3: Timeline of the VR experience in the proposed reference implementation.

pictogram that better represents the way you felt while recalling the previous memory" (see Figure 5.4). The data from the questionnaires, timestamps, markers, and EEG signals are stored for offline analysis.

5.4 The signal processing module

An affect detection process is proposed based on the results obtained in chapter 4 of this thesis. The steps of this process are:

1. **Notch filter:** Remove powerline noise using a notch filter at 50 Hz or 60 Hz, depending on the powerline frequency of the country where the system is used.

2. **Remove bad channels:** Remove channels where a flat-line longer than 5 seconds is detected or whose correlation with nearby channels is lower than 80%.

3. **Remove artifacts:** Remove artifacts caused by eye movements, eye blinks and other noise sources using Artifact Subspace Reconstruction (ASR) [135].

4. **Re-referencing:** Conduct common-average re-referencing.

5. **Remove frequencies:** Apply band-pass filter to remove frequencies below 4 Hz and above 45 Hz.

6. **Down-sampling:** Down-sample to 128 Hz to increase computer processing speed.

7. **Power extraction:** Use Welch's method to extract the power at each power band (delta, theta, alpha, beta and gamma). An 8 seconds Hann window is used. The process is conducted at electrode sites F3, F4, P3 and P4.

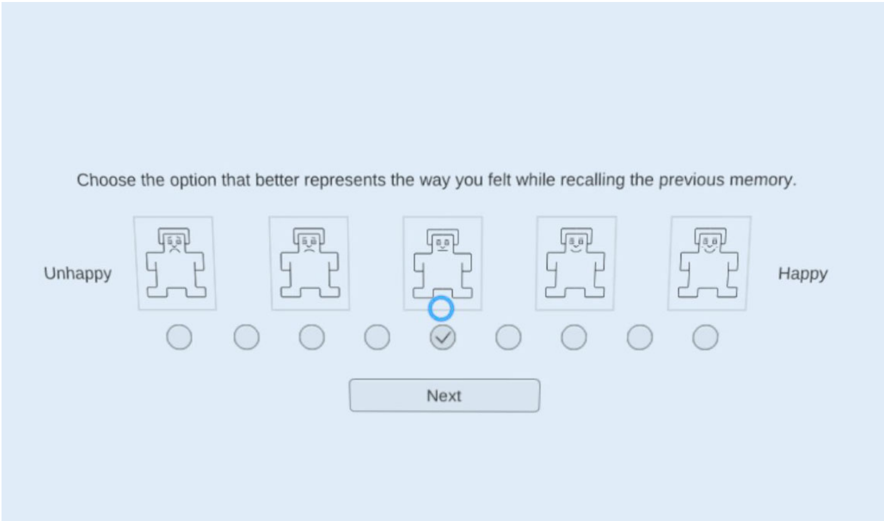


Figure 5.4: Screenshot of the SAM [105] questionnaire implemented inside the virtual environment.

8. **Calculate relative power:** Divide the power in each power band by the total power, yielding the relative power at each power band, at each electrode site during each sample of each epoch.

9. **Valence detection branch:**

   (a) **Extract frontal asymmetry:** Subtract the relative delta and alpha power at F3 from the relative delta and alpha power at F4.

   (b) **Extract parietal gamma power:** Calculate the mean relative gamma power at electrode sites P3 and P4.

   (c) **Build valence model:** Use frontal delta and alpha power, as well as parietal gamma power to train a Linear Discriminant Analysis (LDA) classification model (low valence vs. high valence).

   (d) **Stream valence label:** Send valence labels to the virtual environment using LSL.

10. **Arousal detection branch:**

    (a) **Extract frontal beta asymmetry:** Subtract the relative beta power at F3 from the relative beta power at F4.

    (b) **Train arousal model:** Use frontal beta asymmetry to train a Linear Discriminant Analysis (LDA) classification model for arousal (low arousal vs. high arousal).

    (c) **Stream arousal label:** Send arousal labels to Unity using LSL.

The signal processing module is implemented in NeuroPype (Intheon Labs, California), a software for biosignal processing. There were four main challenges during the implementation of the affect detection technique in this software.

Firstly, the software does not allow to perform down-sampling during online signal processing. Therefore, this step was omitted. Secondly, NeuroPype (Intheon Labs, California) stopped streaming the classification labels to Unity at random intervals. The author of this thesis failed to identify the underlying cause for this issue.

Thirdly, it was not possible to obtain an output from the LDA node when Welch's method was used to extract the spectral content of the signal. The workaround for this problem was to use an FIR filter instead of the Welch's method. It is not clear why the pipeline worked with the FIR filter instead of Welch's method. This problem occurred in NeuroPype (Intheon Labs, California) but not in the previous tests conducted in Python (see chapter 4).

Fourthly, the proposed technique consisted of a classification model for valence and another for arousal. The aim was to implement both classification models as two branches of the same pipeline. However, it was not possible to stream the output of both classification models when they were built in the same NeuroPype (Intheon Labs, California) pipeline. The workaround for this challenge was to create two separate Neuropype (Intheon Labs, California) pipelines: one for valence and another for arousal. Each pipeline is run in a separate Neuropype (Intheon Labs, California) window. Screenshots of the valence and arousal pipelines are presented in Figures 5.5 and 5.6, respectively.

5.5 Discussion

Sentui is a reference implementation of a VR neurofeedback tool for training affective self-regulation. The system consists of two modules. The communication between both modules is established through the WiFi network, using LabStreamingLayer (LSL) [186].

Affective states are triggered with the autobiographical memories of the user, while brain activity is monitored with EEG sensors. The EEG signals are analyzed in real-time using an affect detection technique implemented in Neuropype (Intheon Labs, California). The information extracted from the EEG signals is used to train two support vector machine models: one for valence and another for arousal. The output of both models is used to adapt the VR stimuli, providing visual feedback to the users about their affective states. It is expected that continued use of the system would increase the ability of users to modulate their affective states [44].

The VR stimuli consist of four virtual environments previously tested in an experiment with human participants (see chapter 3). According to the results obtained during that experiment, each of the four virtual environments tends to be associated with one of the four quadrants of the CMA [65]. The transition between the virtual environments is controlled by simple if/then adaptive rules. Future studies could analyze how to build more complex rules for adaptivity. For example, creating personalized visualizations based on the individual characteristics of each user. Previous studies have shown that it is possible to create visual compositions automatically in VR [99]. Furthermore, other studies have shown that it is possible to create content automatically in video games according to the preferences of the user [96, 97]. Therefore, it might be possible to identify the visual cues that each user tends to associate with an affective state and create personalized VR content accordingly [33].

The user experience could be improved by reducing the required time to train the affect detection model. This could be achieved in at least two ways. One possibility is storing the EEG data of the user and training the model with the stored data. This approach would allow training the model only during the first time the system is used, and the accuracy would improve over time as more user data is stored. Another possibility is building an user-independent affect detection model. That is, a model trained with data from a standard dataset, avoiding the requirement of having a calibration phase. For instance, Semertzidis et al. [21] developed an AR system integrated into a BCI for affect communication among dyads. Their affect detection model was trained with data from the DEAP dataset [23], using a procedure proposed by Liu et al. [187]. The accuracy of their model was 58.33 %. Additional studies could implement the affect detection technique proposed in chapter 4 with software different from NeuroPype. For example, using OpenViVE [188], MNE-Python [189] or BCILAB [137].

The primary purpose of Sentui is training users' ability to regulate their affective states voluntarily. However, its basic structure consists of a closed affect loop that could be used for other applications. For example, to develop adaptive VR systems where the input is the user's affective state. Furthermore, it might be used for estimating users' satisfaction with a VR system, similar to Antons [164]. Yet, there are challenges related to the implementation of the affect detection technique in a brain-computer interfacing software that limit the stability of the system. One possibility to solve those challenges consists on implementing the proposed affect detection technique in a software similar to NeuroPype (Intheon Labs, California), such as OpenVive [188], BCILab [137] or MNE-Python [189].

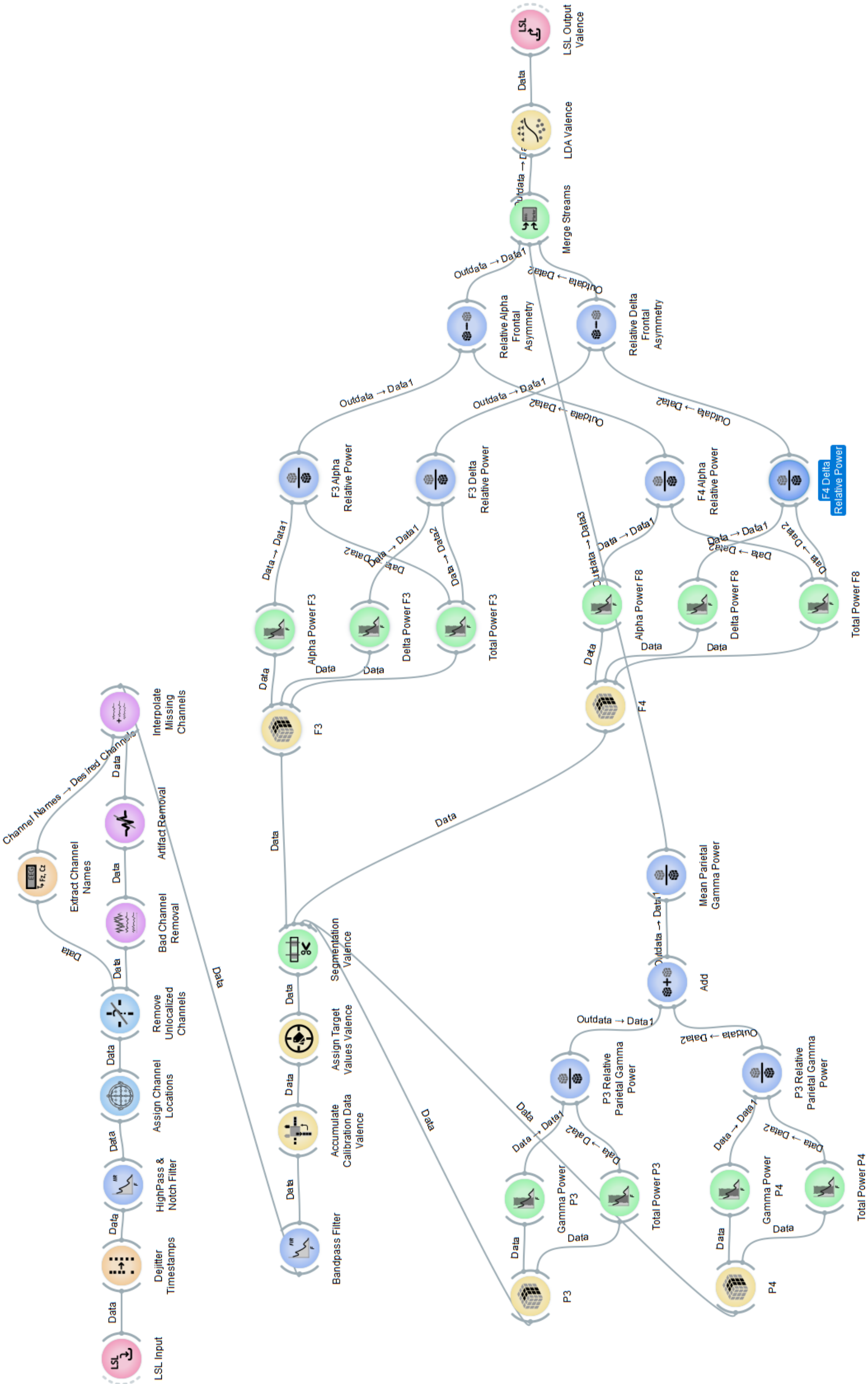


Figure 5.5: Screenshot of the valence pipeline implemented in NeuroPype (Intheon Labs, California).

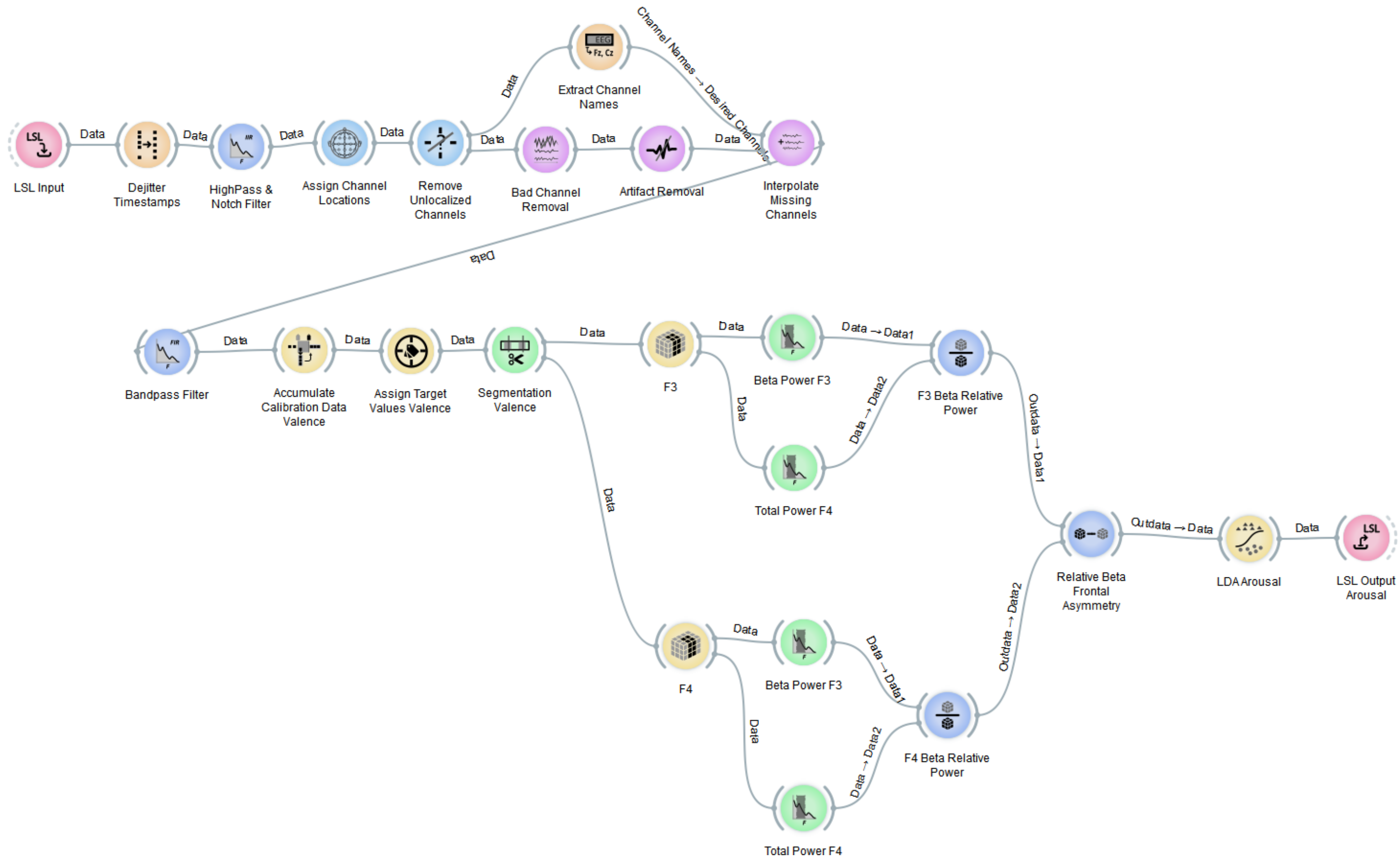


Figure 5.6: Screenshot of the arousal pipeline implemented in NeuroPype (Intheon Labs, California).

6
General discussion

This thesis explores how to develop a VR system that automatically adjusts its content according to the user's affective state. This was achieved by developing a reference implementation of an affective neurofeedback VR system. This system infers users' affective states from their brain activity. Then, the detected affective states are used to execute if/then commands to adjust the VR content. Therefore, the architecture of the system resembles a closed feedback loop (see Figure 6.1).

The proposed reference implementation consists of two modules: the VR content module and the signal processing module. The VR content module consists of four virtual environments that were developed following design guidelines identified in previous studies (see Table 2.1) and validated in an experiment with human participants (see section 3.4) . The results obtained in this experiment suggested that the virtual environments tend to represent four different affective states, regardless of individual differences across participants. Moreover, each virtual environment is associated with one of the four quadrants of the Circumplex Model of Affect (CMA) [65]. Therefore, those four virtual environments were used to develop the neurofeedback VR system described in chapter 5 (see section 5.3).

The signal processing module is based on an affect detection method tested with data taken from the DEAP dataset [23] (see sections 4.1 and 4.2). The test was conducted in an offline setting, emulating the steps that would be conducted in an online analysis. Therefore, the proposed affect detection method was implemented in NeuroPype (Intheon Labs, California), a platform for brain-computer interfacing (see section 5.4). Two models are trained in NeuroPype: one for valence and another for arousal. The models are trained with data recorded from the user during an initial calibration phase.

A pilot experiment with human participants was conducted to explore the implementation of the proposed affect detection method in a VR system (see section 4.3). In this experiment, participants evaluated sixteen videos of the DEAP dataset [23] in terms of the three dimensions of the Evaluative Space Model (ESM) [66]: negativity, positivity and net predisposition. The results do not indicate a clear advantage in using a three-dimensional model instead of a two-dimensional model (see Appendix A). Therefore, using three dimensions might add unnecessary complexity to the system. Consequently, the reference implementation proposed in chapter 5 is based on the two dimensions of the CMA [65]: valence and arousal. Nonetheless, the results obtained in the pilot study suggest that it is possible to use the proposed technique to infer affective states of VR users. Moreover, the results indicate that it is possible to implement the technique using only seven EEG electrodes. Using a small amount of EEG electrodes might be convenient for reducing the cost associated to the equipment and for reducing the time that it takes to setup an experiment.

All things considered, the development of the reference implementation proposed in chapter 5 allows achieving the central purpose of this research project: understanding how to develop a VR system that automatically adjusts its content according to users' affective states. The development of this reference implementation was based on a literature review, two experiments with human participants and offline tests conducted with a standard dataset for emotion research [23].

6.1 Research questions

6.1.1 Visual representation of affective states

1. How to build a virtual environment for affective visualization? The evidence analysed in the literature review (see chapter 2) suggests that there are psychophysiological mechanisms associated with the processing of emotional stimuli that are stable across users [63, 124, 60, 130, 61]. Additional studies indicate that there are visual cues that tend to be associated with similar affective states, regardless of individual differences between users [73, 72, 71, 73, 72, 74]. Based on those studies, four virtual environments were designed. These virtual environments were tested in an experiment where participants evaluated the affective states that each virtual environment was associated with. The results indicate that these virtual environments allow to represent affective states in terms of the four quadrants of the CMA [65]. These virtual environments were designed based on the guidelines summarised in Table 2.1. These guidelines might be useful for further development of visual representations of affective states.

6.1.2 Affect detection technique

2. How to develop a technique for near real-time affect detection that can be implemented in VR systems? The affect detection technique was tested with two versions of the same dataset: the preprocessed and raw versions of the DEAP dataset [23]. Additionally, it was tested with data collected during an experiment conducted by the author of this thesis. The results led to a signal processing method suitable for near real-time analysis of VR users' affective states. This method consists of three main phases: (1) preprocessing, (2) feature extraction, and (3) training of valence and arousal models. The steps of the proposed method are:

1. **Notch filter:** Remove powerline noise using a notch filter at 50 Hz or 60 Hz, depending on the powerline frequency of the country where the system is used.
2. **Remove bad channels:** Remove channels where a flat-line longer than 5 seconds is detected or whose correlation with nearby channels is lower than 80%.
3. **Remove artifacts:** Remove artifacts caused by eye movements, eye blinks and other noise sources using Artifact Subspace Reconstruction (ASR) [135].
4. **Re-referencing:** Conduct common-average re-referencing.

5. **Remove frequencies:** Apply band-pass filter to remove frequencies below 4 Hz and above 45 Hz.

6. **Down-sampling:** Down-sample to 128 Hz to increase computer processing speed.

7. **Power extraction:** Use Welch's method to extract the power at each power band (delta, theta, alpha, beta and gamma). An 8 seconds Hann window is used. The process is conducted at electrode sites F3, F4, P3 and P4.

8. **Calculate relative power:** Divide the power in each power band by the total power, yielding the relative power at each power band, at each electrode site during each sample of each epoch.

9. **Valence detection branch:**

   (a) **Extract frontal asymmetry:** Subtract the relative delta and alpha power at F3 from the relative delta and alpha power at F4.

   (b) **Extract parietal gamma power:** Calculate the mean relative gamma power at electrode sites P3 and P4.

   (c) **Build valence model:** Use frontal delta and alpha power, as well as parietal gamma power to train a Linear Discriminant Analysis (LDA) classification model (low valence vs. high valence).

   (d) **Stream valence label:** Send valence labels to the virtual environment using LSL.

10. **Arousal detection branch:**

   (a) **Extract frontal beta asymmetry:** Subtract the relative beta power at F3 from the relative beta power at F4.

   (b) **Train arousal model:** Use frontal beta asymmetry to train a Linear Discriminant Analysis (LDA) classification model for arousal (low arousal vs. high arousal).

   (c) **Stream arousal label:** Send arousal labels to Unity using LSL.

6.1.3 Closed affect loop in VR systems

3. How to integrate the virtual environment with the affect detection technique? Developing a VR system that can adjust automatically to users' affective states requires building a closed affect loop. That is, a system where users' affective states are instantaneously analyzed, and the outcome of that analysis is used as a computer command. Therefore, the system architecture consists of two modules: the VR content and signal processing modules. The information between both modules is transmitted over the WiFi network using the LabStreamingLayer (LSL) protocol. The architecture of the system is described in Figure 6.1.

6.2 Limitations

An additional experiment was planned to analyze whether the usage of the proposed VR system would lead to increased affective self-regulation. This experiment was canceled due to COVID-19

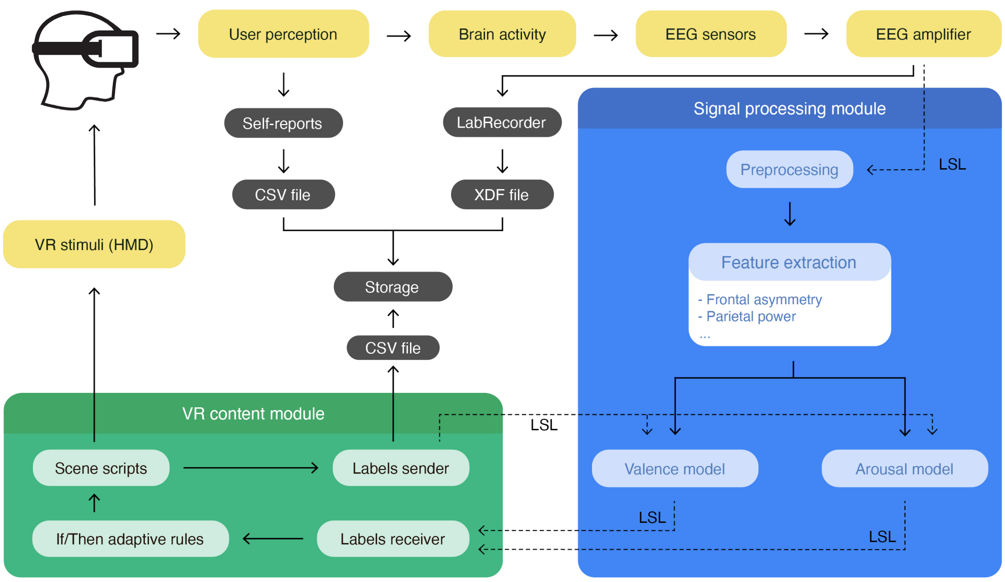


Figure 6.1: Architecture of the proposed reference implementation. The system consists of two modules: the VR content module and the signal processing module. The connection between both modules is established using the LabStreamingLayer (LSL) protocol [186].

restrictions. Therefore, it is not possible to analyze whether the system is effective for training users' ability to modulate their affective states.

The affect detection technique proposed in this thesis is based on previous studies indicating that frontal asymmetry and parietal activity are associated with affective responses [61, 126, 127, 123, 23]. However, many other features could be extracted from EEG signals to build a more robust technique. For example, Val Calvo [190] proposed a method for analyzing affective responses in real-time. The features are extracted using Differential Entropy, Amplitude Envelope, Petrosian Fractal Dimension, Higuchi Fractal Dimension, and Fisher Information. Then, features are selected using the chi-squared statistic. Likewise, Martinez-Tejada et al. [70] proposed to extract time-domain and frequency-domain features to estimate emotional reactions of video game players.

Overall, the performance metrics obtained with the technique were better for valence than for arousal (see Tables 4.2, 4.4). Given that Machine Learning models tend to perform better as more data is available, it is reasonable to assume that the performance of the models would increase if the technique is applied in a dataset containing more trials per participant. Besides, a largest dataset could lead to more robust performance metrics because two of the methods used for noise reduction (ASR for artifact removal [135] and common average re-referencing) tend to perform better as more data is available.

6.3 Future work

The adaptivity of the proposed system consists of basic if/then commands. It is possible to include an additional module in the reference implementation for implementing more complex adaptive rules. The role of this module would be to decide what to adapt and how to adapt it, similar to Bermudez i Badia [33].

Previous studies indicate that emotions tend to spread across groups of people automatically [182, 191], as a consequence of a psychological phenomenon known as emotional contagion [192]. At the same time, it has been found that negative emotions tend to spread more easily than positive emotions [66, 149]. Future studies could analyze whether providing affective feedback to crowds of VR users can help regulate the spread of negative emotions. This might help reduce the contagion of negative emotions, improving users' overall experience.

There are multiple potential applications for systems where the user's affective state is monitored. However, there are ethical concerns related to user privacy that must be considered. For example, Facebook conducted a study where the affective state of more than 600.000 users was influenced by manipulating the emotional content of the posts that each user saw on their home page [191]. The development of wearable technologies might lead to massive adoption of bodily sensors, allowing data harvesting companies to collect an enormous amount of information related to the users' mental states. Therefore, there is a risk that this information is used without the awareness of users.

The development of Distributed Ledger Technologies (DLTs), such as blockchain [193] and IOTA [194], might help prevent this risk. On the one hand, blockchain can be used to build automated access-control managers, preventing unauthorized parties from accessing users' private data [193]. On the other hand, IOTA can be used to facilitate interactions between users' devices and third parties without revealing private content stored in the data [195]. Therefore, DLTs might contribute to restrict unauthorized access to private information, such as users' affective states. Thus, these technologies might allow the integration of physiological sensors into VR systems without compromising user privacy.

References

[1] World Health Organization. *Mental health.* URL: https://www.who.int/westernpacific/health-topics/mental-health (visited on 08/15/2021).

[2] Spencer L James et al. "Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017". In: *The Lancet* 392.10159 (Nov. 2018), pp. 1789–1858. ISSN: 01406736. DOI: 10.1016/S0140-6736(18)32279-7.

[3] United Nations, Department of Economic and Social Affairs, and Population Division. *World population ageing, 2019 highlights.* OCLC: 1134853628. 2020. ISBN: 978-92-1-148325-3.

[4] World Health Organization. *The WHO special initiative for mental health (2019-2023): universal health coverage for mental health.* Tech. rep. WHO/MSD/19.1. World Health Organization, 2019.

[5] *COVID-19 disrupting mental health services in most countries, WHO survey.* URL: https://www.who.int/news/item/05-10-2020-covid-19-disrupting-mental-health-services-in-most-countries-who-survey (visited on 08/15/2021).

[6] Martin Jensen Mækelæ, Niv Reggev, Natalia Dutra, Ricardo M. Tamayo, Reinaldo A. Silva-Sobrinho, Kristoffer Klevjer, and Gerit Pfuhl. "Perceived efficacy of COVID-19 restrictions, reactions and their impact on mental health during the early phase of the outbreak in six countries". In: *Royal Society Open Science* 7.8 (Aug. 2020). ISSN: 2054-5703. DOI: 10.1098/rsos.200644.

[7] Hans Joachim Salize, Wulf Rössler, and Thomas Becker. "Mental health care in Germany: Current state and trends". In: *European Archives of Psychiatry and Clinical Neuroscience* 257.2 (Mar. 2007), pp. 92–103. ISSN: 0940-1334, 1433-8491. DOI: 10.1007/s00406-006-0696-9.

[8] Stephen Carbone, Debra Rickwood, and Chris Tanti. "Workforce shortages and their impact on Australian youth mental health service reform". In: *Advances in Mental Health* 10.1 (Oct. 2011), pp. 92–97. ISSN: 1838-7357, 1837-4905. DOI: 10.5172/jamh.2011.10.1.92.

[9] D.H. Barlow. "The nature of anxiety: Anxiety, depression and emotional disorders." In: *R. M. Rapee D. H. Barlow (Eds.), Chronic anxiety: Generalized anxiety disorder and mixed anxiety-depression.* Guilford Press, 1991, pp. 1–28.

[10] Michael Brook and David S. Kosson. "Impaired cognitive empathy in criminal psychopathy: Evidence from a laboratory measure of empathic accuracy". In: *Journal of Abnormal Psychology* 122.1 (2013). Publisher: American Psychological Association, pp. 156–166. ISSN: 1939-1846(Electronic), 0021-843X(Print). DOI: 10.1037/a0030261.

[11] Youssef Shiban, Jonas Reichenberger, Inga D. Neumann, and Andreas Mühlberger. “Social conditioning and extinction paradigm: a translational study in virtual reality”. In: *Frontiers in Psychology* 6 (2015). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2015.00400.

[12] Andrew Perkis, Christian Timmerer, Sabina Baraković, Jasmina Baraković Husić, Søren Bech, Sebastian Bosse, Jean Botev, Kjell Brunnström, Luis Cruz, Katrien De Moor, Andrea de Polo Saibanti, Wouter Durnez, Sebastian Egger-Lampl, Ulrich Engelke, Tiago H. Falk, Asim Hameed, Andrew Hines, Tanja Kojic, Dragan Kukolj, Eirini Liotou, Dragorad Milovanovic, Sebastian Möller, Niall Murray, Babak Naderi, Manuela Pereira, Stuart Perry, Antonio Pinheiro, Andres Pinilla, Alexander Raake, Sarvesh Rajesh Agrawal, Ulrich Reiter, Rafael Rodrigues, Raimund Schatz, Peter Schelkens, Steven Schmidt, Saeed Shafiee Sabet, Ashutosh Singla, Lea Skorin-Kapov, Mirko Suznjevic, Stefan Uhrig, Sara Vlahović, Jan-Niklas Voigt-Antons, and Saman Zadtootaghaj. *QUALINET White Paper on Definitions of Immersive Media Experience (IMEx)*. May 2020.

[13] Valentijn T. Visch, Ed S. Tan, and Dylan Molenaar. “The emotional and cognitive effect of immersion in film viewing”. In: *Cognition & Emotion* 24.8 (2010), pp. 1439–1445. ISSN: 0269-9931, 1464-0600. DOI: 10.1080/02699930903498186.

[14] Steven Schmidt, Patrick Ehrenbrink, Benjamin Weiss, Jan-Niklas Voigt-Antons, Tanja Kojic, Andrew Johnston, and Sebastian Möller. “Impact of Virtual Environments on Motivation and Engagement During Exergames”. In: *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. Cagliari: IEEE, 2018, pp. 1–6. ISBN: 978-1-5386-2605-4. DOI: 10.1109/QoMEX.2018.8463389.

[15] Thomas K. Metzinger. “Why Is Virtual Reality Interesting for Philosophers?” In: *Frontiers in Robotics and AI* 5 (2018), p. 101. ISSN: 2296-9144. DOI: 10.3389/frobt.2018.00101.

[16] Jaime A Garcia and Karla Felix Navarro. “StepKinnection: a fall prevention game mindfully designed for the elderly”. In: *Studies in health technology and informatics* 214 (2015), pp. 43–49.

[17] D. Z. Blandón, J. E. Muñoz, D. S. Lopez, and O. H. Gallo. “Influence of a BCI neurofeedback videogame in children with ADHD. Quantifying the brain activity through an EEG signal processing dedicated toolbox”. In: *2016 IEEE 11th Colombian Computing Conference (CCC)*. 2016, pp. 1–8. DOI: 10.1109/ColumbianCC.2016.7750788.

[18] Johannes Blum, Christoph Rockstroh, and Anja S. Göritz. “Development and Pilot Test of a Virtual Reality Respiratory Biofeedback Approach”. In: *Applied Psychophysiology and Biofeedback* 45.3 (2020), pp. 153–163. ISSN: 1090-0586, 1573-3270. DOI: 10.1007/s10484-020-09468-x. URL: https://link.springer.com/10.1007/s10484-020-09468-x.

[19] G. N. Yannakakis and Ana Paiva. “Emotion in games”. In: *Handbook on affective computing* (2014). Publisher: Oxford University Press, pp. 459–471.

[20] Thorsten O Zander and Christian Kothe. “Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general”. In: *Journal of Neural Engineering* 8.2 (Apr. 2011), p. 025005. ISSN: 1741-2560, 1741-2552. DOI: 10.1088/1741-2560/8/2/025005.

[21] Nathan Semertzidis, Michaela Scary, Josh Andres, Brahmi Dwivedi, Yutika Chandrashekhar Kulwe, Fabio Zambetta, and Florian Floyd Mueller. "Neo-Noumena: Augmenting Emotion Communication". In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, 2020, pp. 1–13. ISBN: 978-1-4503-6708-0. DOI: 10.1145/3313831.3376599.

[22] Jon H. Kaas. "The evolution of brains from early mammals to humans: Evolution of brains from early mammals to humans". In: *Wiley Interdisciplinary Reviews: Cognitive Science* 4.1 (2013), pp. 33–45. ISSN: 19395078. DOI: 10.1002/wcs.1206.

[23] S. Koelstra, C. Muhl, M. Soleymani, Jong-Seok Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras. "DEAP: A Database for Emotion Analysis ;Using Physiological Signals". In: *IEEE Transactions on Affective Computing* 3.1 (2012), pp. 18–31. ISSN: 1949-3045. DOI: 10.1109/T-AFFC.2011.15.

[24] Andres Pinilla, Jaime Garcia, William Raffe, Jan-Niklas Voigt-Antons, Robert P. Spang, and Sebastian Möller. "Affective Visualization in Virtual Reality: An Integrative Review". In: *Frontiers in Virtual Reality* 2 (2021), p. 630731. ISSN: 2673-4192. DOI: 10.3389/frvir.2021.630731.

[25] Jaime Andres Garcia and Karla Felix Navarro. "The Mobile RehApp™: an AR-based mobile game for ankle sprain rehabilitation". In: *2014 IEEE 3nd International Conference on Serious Games and Applications for Health (SeGAH)*. Rio de Janeiro: IEEE, 2014, pp. 1–6. ISBN: 978-1-4799-4823-9. DOI: 10.1109/SeGAH.2014.7067087.

[26] Sebastian Arndt, Andrew Perkis, and Jan-Niklas Voigt-Antons. "Using Virtual Reality and Head-Mounted Displays to Increase Performance in Rowing Workouts". In: *Proceedings of the 1st International Workshop on Multimedia Content Analysis in Sports - MMSports'18*. Seoul, Republic of Korea: ACM Press, 2018, pp. 45–50. ISBN: 978-1-4503-5981-8. DOI: 10.1145/3265845.3265848.

[27] Robert Greinacher, Tanja Kojic, Luis Meier, Rudresha Gulaganjihalli Parameshappa, Sebastian Moller, and Jan-Niklas Voigt-Antons. "Impact of Tactile and Visual Feedback on Breathing Rhythm and User Experience in VR Exergaming". In: *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. Athlone, Ireland: IEEE, 2020, pp. 1–6. ISBN: 978-1-72815-965-2. DOI: 10.1109/QoMEX48832.2020.9123141.

[28] Grigore C. Burdea, Daniel Cioi, Angad Kale, William E. Janes, Sandy A. Ross, and Jack R. Engsberg. "Robotics and Gaming to Improve Ankle Strength, Motor Control, and Function in Children With Cerebral Palsy—A Case Study Series". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21.2 (2013), pp. 165–173. ISSN: 1534-4320, 1558-0210. DOI: 10.1109/TNSRE.2012.2206055.

[29] J. Hernandez, Y. Li, J. M. Rehg, and R. W. Picard. "BioGlass: Physiological parameter estimation using a head-mounted wearable device". In: *2014 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*. 2014, pp. 55–58. DOI: 10.1109/MOBIHEALTH.2014.7015908.

[30] Grace Leslie, Rosalind Picard, and Simon Lui. “An EEG and Motion Capture Based Expressive Music Interface for Affective Neurofeedback”. In: *Conference: The 1st International Workshop on Brain-Computer Interfacing*. 2015. DOI: 10.13140/RG.2.1.4378.6081.

[31] R.W. Picard, E. Vyzas, and J. Healey. “Toward machine emotional intelligence: analysis of affective physiological state”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23.10 (2001), pp. 1175–1191. ISSN: 01628828. DOI: 10.1109/34.954607.

[32] Ranganatha Sitaram, Sangkyun Lee, Sergio Ruiz, Mohit Rana, Ralf Veit, and Niels Birbaumer. “Real-time support vector classification and feedback of multiple emotional brain states”. In: *NeuroImage* 56.2 (2011), pp. 753–765. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.08.007.

[33] Sergi Bermudez i Badia, Luis Velez Quintero, Monica S. Cameirao, Alice Chirico, Stefano Triberti, Pietro Cipresso, and Andrea Gaggioli. “Toward Emotionally Adaptive Virtual Reality for Mental Health Applications”. In: *IEEE Journal of Biomedical and Health Informatics* 23.5 (2019), pp. 1877–1887. ISSN: 2168-2194, 2168-2208. DOI: 10.1109/JBHI.2018.2878846.

[34] Sebastian Koenig, Gregory Crucian, Andreas Duenser, Christoph Bartneck, and John Dalrymple-Alford. “Validity evaluation of a spatial memory task in virtual environments”. In: *International Journal of Design and Innovation Research* (2011).

[35] Julia Belger, Angelika Thone-Otto, Stephan Krohn, Carsten Finke, Johanne Tromp, Felix Klotzsche, Arno Villringer, Michael Gaebler, Paul Chojecki, and Eva Quinque. “Immersive Virtual Reality for the Assessment and Training of Spatial Memory: Feasibility in Individuals with Brain Injury”. In: *2019 International Conference on Virtual Rehabilitation (ICVR)*. Tel Aviv, Israel: IEEE, 2019, pp. 1–2. ISBN: 978-1-72811-285-5. DOI: 10.1109/ICVR46560.2019.8994342.

[36] Johannes Blum, Christoph Rockstroh, and Anja S. Göritz. “Heart Rate Variability Biofeedback Based on Slow-Paced Breathing With Immersive Virtual Reality Nature Scenery”. In: *Frontiers in Psychology* 10 (2019), p. 2172. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2019.02172.

[37] Anna Felnhofer, Oswald D. Kothgassner, Mareike Schmidt, Anna-Katharina Heinzle, Leon Beutl, Helmut Hlavacs, and Ilse Kryspin-Exner. “Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios”. In: *International Journal of Human - Computer Studies* 82.Complete (2015), pp. 48–56. DOI: 10.1016/j.ijhcs.2015.05.004.

[38] Nilufar Baghaei, Lehan Stemmet, Imran Khaliq, Atefeh Ahmadi, Ilona Halim, Hai-Ning Liang, Wenge Xu, Mark Billinghurst, and Richard Porter. “Designing Individualised Virtual Reality Applications for Supporting Depression: A Feasibility Study”. In: *Companion of the 2021 ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. EICS ’21. Virtual Event, Netherlands: Association for Computing Machinery, 2021, pp. 6–11. ISBN: 9781450384490. DOI: 10.1145/3459926.3464761. URL: https://doi.org/10.1145/3459926.3464761.

[39] Youssef Shiban, Henrik Peperkorn, Georg W. Alpers, Paul Pauli, and Andreas Mühlberger. “Influence of perceptual cues and conceptual information on the activation and reduction of claustrophobic fear”. In: *Journal of Behavior Therapy and Experimental Psychiatry* 51 (2016), pp. 19–26. ISSN: 00057916. DOI: 10.1016/j.jbtep.2015.11.002.

[40] Henrik M. Peperkorn, Georg W. Alpers, and Andreas Mühlberger. "Triggers of Fear: Perceptual Cues Versus Conceptual Information in Spider Phobia: Cues Versus Information in Spider Phobia". In: *Journal of Clinical Psychology* 70.7 (2014), pp. 704–714. ISSN: 00219762. DOI: 10.1002/jclp.22057.

[41] Marc Cavazza, Gabor Aranyi, Fred Charles, Julie Porteous, Stephen Gilroy, Ilana Klovatch, Gilan Jackont, Eyal Soreq, Nimrod Jakob Keynan, Avihay Cohen, Gal Raz, and Talma Hendler. "Towards Empathic Neurofeedback for Interactive Storytelling". In: (2014). Publisher: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik GmbH, Wadern/Saarbruecken, Germany, 19 pages. DOI: 10.4230/OASICS.CMN.2014.42.

[42] Desmet. "Design for mood: Twenty activity-based opportunities to design for mood regulation". In: *International Journal of Design* 9.2 (2015).

[43] Kunal Gupta, Sam W. T. Chan, Yun Suen Pai, Alexander Sumich, Suranga Nanayakkara, and Mark Billinghurst. "Towards Understanding Physiological Responses to Emotional Autobiographical Memory Recall in Mobile VR Scenarios". In: *Adjunct Publication of the 23rd International Conference on Mobile Human-Computer Interaction.* New York, NY, USA: Association for Computing Machinery, 2021. ISBN: 9781450383295. URL: https://doi.org/10.1145/3447527.3474864.

[44] Zhonglin Li, Li Tong, Linyuan Wang, Yongli Li, Wenjie He, Min Guan, and Bin Yan. "Self-regulating positive emotion networks by feedback of multiple emotional brain states using real-time fMRI". In: *Experimental Brain Research* 234.12 (Dec. 2016), pp. 3575–3586. ISSN: 0014-4819, 1432-1106. DOI: 10.1007/s00221-016-4744-z.

[45] Hilary Putnam. "The nature of mental states". In: *Art, Mind, and Religion.* Ed. by W. H. Capitan and D. D. Merrill. Pittsburgh University Press, 1967, pp. 1–223.

[46] Leda Cosmides and John Tooby. "Origins of domain specificity: The evolution of functional organization". In: *Mapping the Mind.* Ed. by Lawrence A. Hirschfeld and Susan A. Gelman. 1st ed. Cambridge University Press, Apr. 1994, pp. 85–116. ISBN: 978-0-521-41966-6 978-0-521-42993-1 978-0-511-75290-2. DOI: 10.1017/CBO9780511752902.005.

[47] Wilhelm Wundt. *Outlines of psychology.* Trans. by Charles Hubbard Judd. London: Williams and Norgate, 1897. DOI: 10.1037/12908-000.

[48] Lisa Feldman Barrett and Eliza Bliss-Moreau. "Chapter 4 Affect as a Psychological Primitive". In: *Advances in Experimental Social Psychology.* Vol. 41. Elsevier, 2009, pp. 167–218. ISBN: 978-0-12-374472-2. DOI: 10.1016/S0065-2601(08)00404-8.

[49] Andrew Ortony, Gerald L Clore, and Allan Collins. *The cognitive structure of emotions.* Cambridge university press, 1988.

[50] Cristina Conati and Xiaoming Zhou. "Modeling Students' Emotions from Cognitive Appraisal in Educational Games". In: *Intelligent Tutoring Systems.* Ed. by Stefano A. Cerri, Guy Gouardères, and Fàbio Paraguaçu. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 944–954. ISBN: 978-3-540-47987-1.

[51] Patricia Augustin Jaques and Rosa Maria Vicari. "A BDI approach to infer student's emotions in an intelligent learning environment". In: *Computers & Education* 49.2 (2007), pp. 360–384. ISSN: 03601315. DOI: 10.1016/j.compedu.2005.09.002.

[52] Janani Harischandra and M.U.S Perera. “Intelligent emotion recognition system using brain signals (EEG)”. In: *2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences*. Langkawi, Malaysia: IEEE, Dec. 2012, pp. 454–459. ISBN: 978-1-4673-1666-8 978-1-4673-1664-4. DOI: 10.1109/IECBES.2012.6498050.

[53] Ifigeneia Mavridou, James T. McGhee, Mahyar Hamedi, Mohsen Fatoorechi, Andrew Cleal, Emili Ballaguer-Balester, Ellen Seiss, Graeme Cox, and Charles Nduka. “FACETEQ interface demo for emotion expression in VR”. In: *2017 IEEE Virtual Reality (VR)*. Los Angeles, CA, USA: IEEE, 2017, pp. 441–442. ISBN: 978-1-5090-6647-6. DOI: 10.1109/VR.2017.7892369.

[54] Robert Plutchik. “A psychoevolutionary theory of emotions”. In: *Social Science Information* 21.4-5 (1982), pp. 529–553. ISSN: 0539-0184, 1461-7412. DOI: 10.1177/053901882021004003.

[55] Carl Georg Lange and William James, eds. *The emotions, Vol. 1*. Baltimore: Williams & Wilkins Co, 1922. DOI: 10.1037/10735-000.

[56] Walter B. Cannon. “The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory”. In: *The American Journal of Psychology* 39.1/4 (1927), p. 106. ISSN: 00029556. DOI: 10.2307/1415404.

[57] Philip Bard. “On emotional expression after decortication with some remarks on certain theoretical views: Part I.” In: *Psychological Review* 41.4 (1934), pp. 309–329. ISSN: 1939-1471, 0033-295X. DOI: 10.1037/h0070765.

[58] Stanley Schachter and Jerome Singer. “Cognitive, social, and physiological determinants of emotional state.” In: *Psychological Review* 69.5 (1962), pp. 379–399. ISSN: 0033-295X. DOI: 10.1037/h0046234.

[59] Michelle N. Shiota and James W. Kalat. *Emotion*. Second Edition. Belmont, CA, USA: Linda Schreiber-Ganster, 2012.

[60] Ulf Dimberg, Monika Thunberg, and Kurt Elmehed. “Unconscious Facial Reactions to Emotional Facial Expressions”. In: *Psychological Science* 11.1 (2000), pp. 86–89. ISSN: 0956-7976, 1467-9280. DOI: 10.1111/1467-9280.00221.

[61] Rene J. Huster, Stephan Stevens, Alexander L. Gerlach, and Fred Rist. “A spectralanalytic approach to emotional responses evoked through picture presentation”. In: *International Journal of Psychophysiology* 72.2 (2009), pp. 212–216. ISSN: 01678760. DOI: 10.1016/j.ijpsycho.2008.12.009.

[62] Julian F. Thayer, Anita L. Hansen, Evelyn Saus-Rose, and Bjorn Helge Johnsen. “Heart Rate Variability, Prefrontal Neural Function, and Cognitive Performance: The Neurovisceral Integration Perspective on Self-regulation, Adaptation, and Health”. In: *Annals of Behavioral Medicine* 37.2 (2009), pp. 141–153. ISSN: 0883-6612, 1532-4796. DOI: 10.1007/s12160-009-9101-z.

[63] Paul Ekman and Wallace V. Friesen. “Constants across cultures in the face and emotion.” In: *Journal of Personality and Social Psychology* 17.2 (1971), pp. 124–129. ISSN: 1939-1315, 0022-3514. DOI: 10.1037/h0030377.

[64] Charles Darwin. *The expression of the emotions in man and animals*. London: John Murray, 1872. DOI: 10.1037/10001-000.

[65] James A. Russell. “A circumplex model of affect.” In: *Journal of Personality and Social Psychology* 39.6 (1980), pp. 1161–1178. ISSN: 0022-3514. DOI: 10.1037/h0077714.

[66] John T. Cacioppo, Wendi L. Gardner, and Gary G. Berntson. “Beyond Bipolar Conceptualizations and Measures: The Case of Attitudes and Evaluative Space”. In: *Personality and Social Psychology Review* 1.1 (1997), pp. 3–25. ISSN: 1088-8683, 1532-7957. DOI: 10.1207/s15327957pspr0101_2.

[67] Alison M. Mattek, George L. Wolford, and Paul J. Whalen. “A Mathematical Model Captures the Structure of Subjective Affect”. In: *Perspectives on Psychological Science* 12.3 (2017), pp. 508–526. ISSN: 1745-6916, 1745-6924. DOI: 10.1177/1745691616685863.

[68] Matthew R. Sutherland and Mara Mather. “Negative arousal amplifies the effects of saliency in short-term memory.” In: *Emotion* 12.6 (2012), pp. 1367–1372. ISSN: 1931-1516, 1528-3542. DOI: 10.1037/a0027860.

[69] Peter J Lang, Margaret M Bradley, Bruce N Cuthbert, et al. “International affective picture system (IAPS): Technical manual and affective ratings”. In: *NIMH Center for the Study of Emotion and Attention* 1.39-58 (1997), p. 3.

[70] Laura Alejandra Martínez-Tejada, Alex Puertas-González, Natsue Yoshimura, and Yasuharu Koike. “Exploring EEG Characteristics to Identify Emotional Reactions under Videogame Scenarios”. In: *Brain Sciences* 11.3 (2021), p. 378. ISSN: 2076-3425. DOI: 10.3390/brainsci11030378.

[71] Moshe Bar and Maital Neta. “Humans Prefer Curved Visual Objects”. In: *Psychological Science* 17.8 (Aug. 2006), pp. 645–648. ISSN: 0956-7976, 1467-9280. DOI: 10.1111/j.1467-9280.2006.01759.x.

[72] Joel Aronoff, Barbara A. Woike, and Lester M. Hyman. “Which are the stimuli in facial displays of anger and happiness? Configurational bases of emotion recognition.” In: *Journal of Personality and Social Psychology* 62.6 (1992), pp. 1050–1066. ISSN: 0022-3514. DOI: 10.1037/0022-3514.62.6.1050.

[73] Joel Aronoff. “How We Recognize Angry and Happy Emotion in People, Places, and Things”. In: *Cross-Cultural Research* 40.1 (2006), pp. 83–105. ISSN: 1069-3971, 1552-3578. DOI: 10.1177/1069397105282597.

[74] J. Farley Norman, Amanda Beers, and Flip Phillips. “Fechner’s Aesthetics Revisited”. In: *Seeing and Perceiving* 23.3 (2010), pp. 263–271. ISSN: 1878-4755, 1878-4763. DOI: 10.1163/187847510X516412.

[75] Chaitanya K. Ryali, Stanny Goffin, Piotr Winkielman, and Angela J. Yu. “From likely to likable: The role of statistical typicality in human social assessment of faces”. In: *Proceedings of the National Academy of Sciences* 117.47 (2020), pp. 29371–29380. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1912343117.

[76] Xiao-Ping Gao, John H. Xin, Tetsuya Sato, Aran Hansuebsai, Marcello Scalzo, Kanji Kajiwara, Shing-Sheng Guan, J. Valldeperas, Manuel José Lis, and Monica Billger. “Analysis of cross-cultural color emotion”. In: *Color Research & Application* 32.3 (2007), pp. 223–229. ISSN: 03612317, 15206378. DOI: 10.1002/col.20321.

[77] Patricia Valdez and Albert Mehrabian. "Effects of color on emotions." In: *Journal of Experimental Psychology: General* 123.4 (1994), pp. 394–409. ISSN: 1939-2222, 0096-3445. DOI: 10.1037/0096-3445.123.4.394.

[78] Lyn Bartram, Abhisekh Patra, and Maureen Stone. "Affective Color in Visualization". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver Colorado USA: ACM, May 2017, pp. 1364–1374. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3026041.

[79] S. E. Palmer and K. B. Schloss. "An ecological valence theory of human color preference". In: *Proceedings of the National Academy of Sciences* 107.19 (May 2010), pp. 8877–8882. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0906172107.

[80] Nilgün Camgöz, Cengiz Yener, and Dilek Güvenç. "Effects of hue, saturation, and brightness on preference". In: *Color Research & Application* 27.3 (2002), pp. 199–207. ISSN: 03612317. DOI: 10.1002/col.10051.

[81] Lisa Wilms and Daniel Oberfeld. "Color and emotion: effects of hue, saturation, and brightness". In: *Psychological Research* 82.5 (Sept. 2018), pp. 896–914. ISSN: 0340-0727, 1430-2772. DOI: 10.1007/s00426-017-0880-8.

[82] Ruby Lipson-Smith, Julie Bernhardt, Edoardo Zamuner, Leonid Churilov, Nick Busietta, and Damian Moratti. "Exploring colour in context using Virtual Reality: Does a room change how you feel?" In: *Virtual Reality* (2020). ISSN: 1359-4338, 1434-9957. DOI: 10.1007/s10055-020-00479-x.

[83] Marcel P. Lucassen, Theo Gevers, and Arjan Gijsenij. "Texture affects color emotion". In: *Color Research & Application* 36.6 (2011), pp. 426–436. ISSN: 03612317. DOI: 10.1002/col.20647.

[84] Yurika Ebe and Hiroyuki Umemuro. "Emotion Evoked by Texture and Application to Emotional Communication". In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. Seoul Republic of Korea: ACM, 2015, pp. 1995–2000. ISBN: 978-1-4503-3146-3. DOI: 10.1145/2702613.2732768.

[85] Chao Feng, Lyn Bartram, and Bernhard E. Riecke. "Evaluating affective features of 3D motionscapes". In: *Proceedings of the ACM Symposium on Applied Perception - SAP '14*. Vancouver, British Columbia, Canada: ACM Press, 2014, pp. 23–30. ISBN: 978-1-4503-3009-1. DOI: 10.1145/2628257.2628264.

[86] Lukasz Piwek, Frank Pollick, and Karin Petrini. "Audiovisual integration of emotional signals from others' social interactions". In: *Frontiers in Psychology* 9 (2015). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2015.00611.

[87] M. Lockyer, L. Bartram, and B. E. Riecke. "Simple Motion Textures for Ambient Affect". In: *Computational Aesthetics in Graphics* Visualization (2011). ISSN: 1816-0859. DOI: 10.2312/COMPAESTH/COMPAESTH11/089-096.

[88] Konstantinos Drossos, Andreas Floros, Andreas Giannakoulopoulos, and Nikolaos Kanellopoulos. "Investigating the Impact of Sound Angular Position on the Listener Affective State". In: *IEEE Transactions on Affective Computing* 6.1 (2015), pp. 27–42. ISSN: 1949-3045. DOI: 10.1109/TAFFC.2015.2392768.

[89] Ana Tajadura-Jiménez, Aleksander Väljamäe, Erkin Asutay, and Daniel Västfjäll. "Embodied auditory perception: The emotional impact of approaching and receding sound sources." In: *Emotion* 10.2 (2010), pp. 216–229. ISSN: 1931-1516, 1528-3542. DOI: 10.1037/a0018422.

[90] Ana Tajadura-Jiménez. *Embodied psychoacoustics: Spatial and multisensory determinants of auditory-induced emotion*. Chalmers University of Technology Gothenburg, 2008.

[91] Ana Tajadura-Jiménez, Pontus Larsson, Aleksander Väljamäe, Daniel Västfjäll, and Mendel Kleiner. "When room size matters: Acoustic influences on emotional responses to sounds." In: *Emotion* 10.3 (June 2010), pp. 416–422. ISSN: 1931-1516, 1528-3542. DOI: 10.1037/a0018423.

[92] Alicia Fernández-Sotos, Antonio Fernández-Caballero, and José M. Latorre. "Influence of Tempo and Rhythmic Unit in Musical Emotion Regulation". In: *Frontiers in Computational Neuroscience* 10 (2016). ISSN: 1662-5188. DOI: 10.3389/fncom.2016.00080.

[93] Gina M. Gerardi and Louann Gerken. "The Development of Affective Responses to Modality and Melodic Contour". In: *Music Perception: An Interdisciplinary Journal* 12.3 (1995), pp. 279–290. ISSN: 07307829, 15338312. DOI: 10.2307/40286184.

[94] Toiviaine Petri. "Exploring relationships between audio features and emotion in music". In: *Frontiers in Human Neuroscience* 3 (2009). ISSN: 1662-5161. DOI: 10.3389/conf.neuro.09.2009.02.033.

[95] Duncan Williams, Alexis Kirke, Eduardo Miranda, Ian Daly, Faustina Hwang, James Weaver, and Slawomir Nasuto. "Affective Calibration of Musical Feature Sets in an Emotionally Intelligent Music Composition System". In: *ACM Transactions on Applied Perception* 14.3 (2017), pp. 1–13. ISSN: 15443558. DOI: 10.1145/3059005.

[96] G. N. Yannakakis and J. Togelius. "Experience-Driven Procedural Content Generation". In: *IEEE Transactions on Affective Computing* 2.3 (2011), pp. 147–161. ISSN: 1949-3045. DOI: 10.1109/T-AFFC.2011.6.

[97] William L. Raffe, Fabio Zambetta, Xiaodong Li, and Kenneth O. Stanley. "Integrated Approach to Personalized Procedural Map Generation Using Evolutionary Algorithms". In: *IEEE Transactions on Computational Intelligence and AI in Games* 7.2 (2015), pp. 139–155. ISSN: 1943-068X, 1943-0698. DOI: 10.1109/TCIAIG.2014.2341665.

[98] Agoston E. Eiben and James E. Smith. *Introduction to evolutionary computing*. Natural computing series. Heidelberg: Springer, 2015. ISBN: 978-3-662-44874-8 978-3-662-44873-1.

[99] Alexandra Kitson, Steve DiPaola, and Bernhard E. Riecke. "Lucid Loop: A Virtual Deep Learning Biofeedback System for Lucid Dreaming Practice". In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow Scotland Uk: ACM, 2019, pp. 1–6. ISBN: 978-1-4503-5971-9. DOI: 10.1145/3290607.3312952.

[100] Theodosis Georgiou and Yiannis Demiris. "Adaptive user modelling in car racing games using behavioural and physiological data". In: *User Modeling and User-Adapted Interaction* 27.2 (2017), pp. 267–311. ISSN: 0924-1868, 1573-1391. DOI: 10.1007/s11257-017-9192-3.

[101] Wenzhuo Zhang, Lin Shu, Xiangmin Xu, and Dan Liao. "Affective Virtual Reality System (AVRS): Design and Ratings of Affective VR Scenes". In: *2017 International Conference on Virtual Reality and Visualization (ICVRV)*. 2017, pp. 311–314. DOI: 10.1109/ICVRV.2017.00072.

[102] Javier Marín-Morales, Juan Luis Higuera-Trujillo, Alberto Greco, Jaime Guixeres, Carmen Llinares, Enzo Pasquale Scilingo, Mariano Alcañiz, and Gaetano Valenza. "Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors". In: *Scientific Reports* 8.1 (Sept. 2018), p. 13657. ISSN: 2045-2322. DOI: 10.1038/s41598-018-32063-4. URL: https://doi.org/10.1038/s41598-018-32063-4.

[103] Robert J. Fisher. "Social Desirability Bias and the Validity of Indirect Questioning". In: *Journal of Consumer Research* 20.2 (1993), p. 303. ISSN: 0093-5301, 1537-5277. DOI: 10.1086/209351.

[104] David Watson, Lee Anna Clark, and Auke Tellegen. "Development and validation of brief measures of positive and negative affect: The PANAS scales." In: *Journal of Personality and Social Psychology* 54.6 (1988), pp. 1063–1070. ISSN: 1939-1315, 0022-3514. DOI: 10.1037/0022-3514.54.6.1063.

[105] 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. ISSN: 00057916. DOI: 10.1016/0005-7916(94)90063-9.

[106] Desmet, Martijn H. Vastenburg, and Natalia Romero. "Mood measurement with Pick-A-Mood: review of current methods and design of a pictorial self-report scale". In: *Journal of Design Research* 14.3 (2016), p. 241. ISSN: 1748-3050, 1569-1551. DOI: 10.1504/JDR.2016.079751.

[107] Isabelle Hupont, Joaquin Gracia, Luis Sanagustin, and Miguel Angel Gracia. "How do new visual immersive systems influence gaming QoE? A use case of serious gaming with Oculus Rift". In: *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*. Pylos-Nestoras: IEEE, 2015, pp. 1–6. ISBN: 978-1-4799-8958-4. DOI: 10.1109/QoMEX.2015.7148110.

[108] Valentin Holzwarth, Johannes Schneider, Joshua Handali, Joy Gisler, Christian Hirt, Andreas Kunz, and Jan vom Brocke. "Towards estimating affective states in Virtual Reality based on behavioral data". In: *Virtual Reality* 25 (Dec. 2021). DOI: 10.1007/s10055-021-00518-1.

[109] Peter Bull. "The interpretation of posture through an alternative methodology to role play". In: *British Journal of Social and Clinical Psychology* 17.1 (1978), pp. 1–6. ISSN: 00071293. DOI: 10.1111/j.2044-8260.1978.tb00888.x.

[110] Patrice Robitaille and Michael J. McGuffin. "Increased affect-arousal in VR can be detected from faster body motion with increased heart rate". In: *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. Montreal Quebec Canada: ACM, 2019, pp. 1–6. ISBN: 978-1-4503-6310-5. DOI: 10.1145/3306131.3317022.

[111] Klaus R. Scherer and James S. Oshinsky. "Cue utilization in emotion attribution from auditory stimuli". In: *Motivation and Emotion* 1.4 (1977), pp. 331–346. ISSN: 0146-7239, 1573-6644. DOI: 10.1007/BF00992539.

[112] Daniel T. Cordaro, Dacher Keltner, Sumjay Tshering, Dorji Wangchuk, and Lisa M. Flynn. "The voice conveys emotion in ten globalized cultures and one remote village in Bhutan." In: *Emotion* 16.1 (2016), pp. 117–128. ISSN: 1931-1516, 1528-3542. DOI: 10.1037/emo0000100.

[113] Asha Kapur, Ajay Kapur, Naznin Virji-Babul, George Tzanetakis, and Peter F. Driessen. "Gesture-Based Affective Computing on Motion Capture Data". In: *Affective Computing and Intelligent Interaction*. Ed. by Jianhua Tao, Tieniu Tan, and Rosalind W. Picard. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 1–7. ISBN: 978-3-540-32273-3.

[114] Thurid Vogt, Elisabeth André, and Nikolaus Bee. "EmoVoice — A Framework for Online Recognition of Emotions from Voice". In: *Perception in Multimodal Dialogue Systems*. Ed. by Elisabeth André, Laila Dybkjær, Wolfgang Minker, Heiko Neumann, Roberto Pieraccini, and Michael Weber. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 188–199. ISBN: 978-3-540-69369-7.

[115] Tim Polzehl, Alexander Schmitt, Florian Metze, and Michael Wagner. "Anger recognition in speech using acoustic and linguistic cues". In: *Speech Communication* 53.9-10 (2011), pp. 1198–1209. ISSN: 01676393. DOI: 10.1016/j.specom.2011.05.002.

[116] Rainer Banse and Klaus R. Scherer. "Acoustic profiles in vocal emotion expression." In: *Journal of Personality and Social Psychology* 70.3 (1996), pp. 614–636. ISSN: 1939-1315, 0022-3514. DOI: 10.1037/0022-3514.70.3.614.

[117] Paul Ekman and Wallace V. Friesen. "Measuring facial movement". In: *Environmental Psychology and Nonverbal Behavior* 1.1 (1976), pp. 56–75. ISSN: 0361-3496, 1573-3653. DOI: 10.1007/BF01115465.

[118] D. McDuff, R. E. Kaliouby, and R. W. Picard. "Crowdsourcing Facial Responses to Online Videos". In: *IEEE Transactions on Affective Computing* 3.4 (2012), pp. 456–468. ISSN: 1949-3045. DOI: 10.1109/T-AFFC.2012.19.

[119] Michael Feffer, Ognjen (Oggi) Rudovic, and Rosalind W. Picard. "A Mixture of Personalized Experts for Human Affect Estimation". In: *Machine Learning and Data Mining in Pattern Recognition*. Ed. by Petra Perner. Cham: Springer International Publishing, 2018, pp. 316–330. ISBN: 978-3-319-96133-0.

[120] Javier Marín-Morales, Juan Luis Higuera-Trujillo, Alberto Greco, Jaime Guixeres, Carmen Llinares, Enzo Pasquale Scilingo, Mariano Alcañiz, and Gaetano Valenza. "Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors". In: *Scientific Reports* 8.1 (Sept. 2018), p. 13657. ISSN: 2045-2322. DOI: 10.1038/s41598-018-32063-4. URL: https://doi.org/10.1038/s41598-018-32063-4.

[121] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, and P. J. Schwartz. "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use". In: *European Heart Journal* 17.3 (1996), pp. 354–381. ISSN: 0195-668X, 1522-9645. DOI: 10.1093/oxfordjournals.eurheartj.a014868.

[122] M. Pagani, F. Lombardi, S. Guzzetti, G. Sandrone, O. Rimoldi, G. Malfatto, S. Cerutti, and A. Malliani. "Power spectral density of heart rate variability as an index of sympatho-vagal interaction in normal and hypertensive subjects". In: *Journal of Hypertension. Supplement: Official Journal of the International Society of Hypertension* 2.3 (Dec. 1984), S383–385. ISSN: 0952-1178.

[123] Simon M. Hofmann, Felix Klotzsche, Alberto Mariola, Vadim V. Nikulin, Arno Villringer, and Michael Gaebler. "Decoding Subjective Emotional Arousal during a Naturalistic VR Experience from EEG Using LSTMs". In: *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. Taichung, Taiwan: IEEE, 2018, pp. 128–131. ISBN: 978-1-5386-9269-1. DOI: 10.1109/AIVR.2018.00026.

[124] Ulf Dimberg. "Facial Reactions to Facial Expressions". In: *Psychophysiology* 19.6 (1982), pp. 643–647. ISSN: 0048-5772, 1469-8986. DOI: 10.1111/j.1469-8986.1982.tb02516.x.

[125] Ulf Dimberg and Monika Thunberg. "Empathy, emotional contagion, and rapid facial reactions to angry and happy facial expressions: Empathy and rapid facial reactions". In: *PsyCh Journal* 1.2 (2012), pp. 118–127. ISSN: 20460252. DOI: 10.1002/pchj.4.

[126] W. Ray and H. Cole. "EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes". In: *Science* 228.4700 (1985), pp. 750–752. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.3992243.

[127] Jan-Niklas Antons, Sebastian Arndt, Robert Schleicher, and Sebastian Möller. "Brain Activity Correlates of Quality of Experience". In: *Quality of Experience*. Ed. by Sebastian Möller and Alexander Raake. Series Title: T-Labs Series in Telecommunication Services. Cham: Springer International Publishing, 2014, pp. 109–119. ISBN: 978-3-319-02680-0 978-3-319-02681-7. DOI: 10.1007/978-3-319-02681-7_8. URL: http://link.springer.com/10.1007/978-3-319-02681-7_8.

[128] G. Pfurtscheller and F.H. Lopes da Silva. "Event-related EEG/MEG synchronization and desynchronization: basic principles". In: *Clinical Neurophysiology* 110.11 (1999), pp. 1842–1857. ISSN: 13882457. DOI: 10.1016/S1388-2457(99)00141-8.

[129] L.I. Aftanas, A.A. Varlamov, S.V. Pavlov, V.P. Makhnev, and N.V. Reva. "Affective picture processing: event-related synchronization within individually defined human theta band is modulated by valence dimension." In: *Neuroscience Letters* 303.2 (May 2001), pp. 115–118. ISSN: 03043940. DOI: 10.1016/S0304-3940(01)01703-7.

[130] Richard J. Davidson. "Emotion and Affective Style: Hemispheric Substrates". In: *Psychological Science* 3.1 (Jan. 1992), pp. 39–43. ISSN: 0956-7976, 1467-9280. DOI: 10.1111/j.1467-9280.1992.tb00254.x.

[131] Boris Reuderink, Christian Mühl, and Mannes Poel. "Valence, arousal and dominance in the EEG during game play". In: *International Journal of Autonomous and Adaptive Communications Systems* 6.1 (2013), p. 45. ISSN: 1754-8632, 1754-8640. DOI: 10.1504/IJAACS.2013.050691.

[132] Marius Klug and Klaus Gramann. *Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments*. preprint. Neuroscience, 2020. DOI: 10.1101/2020.06.02.129213.

[133] S. Makeig, T.-P. Jung, A. J. Bell, D. Ghahremani, and T. J. Sejnowski. “Blind separation of auditory event-related brain responses into independent components”. In: *Proceedings of the National Academy of Sciences* 94.20 (Sept. 1997), pp. 10979–10984. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.94.20.10979.

[134] Luca Pion-Tonachini, Sheng-Hsiou Hsu, Scott Makeig, Tzyy-Ping Jung, and Gert Cauwenberghs. “Real-time EEG Source-mapping Toolbox (REST): Online ICA and source localization”. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Milan: IEEE, 2015, pp. 4114–4117. ISBN: 978-1-4244-9271-8. DOI: 10.1109/EMBC.2015.7319299.

[135] Tim R. Mullen, Christian A. E. Kothe, Yu Mike Chi, Alejandro Ojeda, Trevor Kerth, Scott Makeig, Tzyy-Ping Jung, and Gert Cauwenberghs. “Real-time neuroimaging and cognitive monitoring using wearable dry EEG”. In: *IEEE Transactions on Biomedical Engineering* 62.11 (2015), pp. 2553–2567. ISSN: 0018-9294, 1558-2531. DOI: 10.1109/TBME.2015.2481482.

[136] Sarah Blum, Nadine S. J. Jacobsen, Martin G. Bleichner, and Stefan Debener. “A Riemannian Modification of Artifact Subspace Reconstruction for EEG Artifact Handling”. In: *Frontiers in Human Neuroscience* 13 (2019). ISSN: 1662-5161. DOI: 10.3389/fnhum.2019.00141.

[137] Christian Andreas Kothe and Scott Makeig. “BCILAB: a platform for brain–computer interface development”. In: *Journal of Neural Engineering* 10.5 (Oct. 2013), p. 056014. ISSN: 1741-2560, 1741-2552. DOI: 10.1088/1741-2560/10/5/056014.

[138] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. “Brain–computer interfaces for communication and control”. In: *Clinical Neurophysiology* 113.6 (2002), pp. 767–791. ISSN: 13882457. DOI: 10.1016/S1388-2457(02)00057-3.

[139] Kunal Gupta, Jovana Lazarevic, Yun Suen Pai, and Mark Billinghurst. “AffectivelyVR: Towards VR Personalized Emotion Recognition”. In: *26th ACM Symposium on Virtual Reality Software and Technology*. VRST ’20. Virtual Event, Canada: Association for Computing Machinery, 2020. ISBN: 9781450376198. DOI: 10.1145/3385956.3422122. URL: https://doi.org/10.1145/3385956.3422122.

[140] Raymundo Cassani, Marc-Antoine Moinnereau, and Tiago H. Falk. “A Neurophysiological Sensor-Equipped Head-Mounted Display for Instrumental QoE Assessment of Immersive Multimedia”. In: *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. Cagliari: IEEE, 2018, pp. 1–6. ISBN: 978-1-5386-2605-4. DOI: 10.1109/QoMEX.2018.8463422.

[141] Timofei I. Voznenko, Eugene V. Chepin, and Gleb A. Urvanov. “The Control System Based on Extended BCI for a Robotic Wheelchair”. In: *Procedia Computer Science* 123 (2018), pp. 522–527. ISSN: 18770509. DOI: 10.1016/j.procs.2018.01.079.

[142] Hossam Almahasneh, Weng-Tink Chooi, Nidal Kamel, and Aamir Saeed Malik. “Deep in thought while driving: An EEG study on drivers’ cognitive distraction”. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 26 (2014), pp. 218–226. ISSN: 13698478. DOI: 10.1016/j.trf.2014.08.001.

[143] Elke A. Schoneveld, Monique Malmberg, Anna Lichtwarck-Aschoff, Geert P. Verheijen, Rutger C.M.E. Engels, and Isabela Granic. "A neurofeedback video game (MindLight) to prevent anxiety in children: A randomized controlled trial". In: *Computers in Human Behavior* 63 (2016), pp. 321–333. ISSN: 07475632. DOI: 10.1016/j.chb.2016.05.005.

[144] José Unpingco. *Python for signal processing: featuring IPython notebooks*. OCLC: 862972486. Cham: Springer, 2014. ISBN: 978-3-319-01342-8 978-3-319-01341-1.

[145] Orly Fuhrman and Lera Boroditsky. "Cross-Cultural Differences in Mental Representations of Time: Evidence From an Implicit Nonlinguistic Task". In: *Cognitive Science* 34.8 (2010), pp. 1430–1451. ISSN: 03640213. DOI: 10.1111/j.1551-6709.2010.01105.x.

[146] A. J. H. Haar, A. Jain, F. Schoeller, and P. Maes. "Augmenting aesthetic chills using a wearable prosthesis improves their downstream effects on reward and social cognition". In: *Scientific Reports* 10.1 (Dec. 2020), p. 21603. ISSN: 2045-2322. DOI: 10.1038/s41598-020-77951-w.

[147] F. Schoeller, P. Bertrand, L.J. Gerry, A. Jain, A.H. Horowitz, and F. Zenasni. "Combining Virtual Reality and Biofeedback to Foster Empathic Abilities in Humans". In: *Frontiers in Psychology* 9 (2019), p. 2741. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.02741.

[148] Lorenza S. Colzato, Roberta Sellaro, and Christian Beste. "Darwin revisited: The vagus nerve is a causal element in controlling recognition of other's emotions". In: *Cortex* 92 (2017), pp. 95–102. ISSN: 00109452. DOI: 10.1016/j.cortex.2017.03.017.

[149] Andrés Pinilla, Ricardo M. Tamayo, and Jorge Neira. "How Do Induced Affective States Bias Emotional Contagion to Faces? A Three-Dimensional Model". In: *Frontiers in Psychology* 11 (2020). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2020.00097.

[150] Andres Pinilla, Jaime Garcia, William Raffe, Jan-Niklas Voigt-Antons, and Sebastian Möller. *Visual representation of emotions in Virtual Reality*. preprint. PsyArXiv, 2021. DOI: 10.31234/osf.io/9jguh.

[151] N. Sobell and M. Trivich. "Brainwave Drawing Game". In: *Delicate Balance: Technics, Culture and Consequences*. Los Angeles, CA, USA: IEEE, 1990, pp. 360–362. ISBN: 978-0-7803-0044-6. DOI: 10.1109/TCAC.1989.697094.

[152] Christoph De Boeck. *Staalhemel - Steel Sky*. 2009. URL: https://vimeo.com/7746875.

[153] Raffaella Folgieri, Francesco Soave, and Claudio Lucchiari. "Creative thinking: a brain computer interface of art". In: *International Conference on Live Interfaces (ICLI2016)*. 2016, pp. 74–79. DOI: 10.13140/RG.2.2.31683.40489.

[154] Vadim Zotev, Raquel Phillips, Han Yuan, Masaya Misaki, and Jerzy Bodurka. "Self-regulation of human brain activity using simultaneous real-time fMRI and EEG neurofeedback". In: *NeuroImage* 85 (Jan. 2014), pp. 985–995. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2013.04.126.

[155] Anna Zilverstand, Bettina Sorger, Pegah Sarkheil, and Rainer Goebel. "fMRI neurofeedback facilitates anxiety regulation in females with spider phobia". In: *Frontiers in Behavioral Neuroscience* 9 (June 2015). ISSN: 1662-5153. DOI: 10.3389/fnbeh.2015.00148.

[156] Gary D. James, Lily S. Yee, and Thomas G. Pickering. "Winter-summer differences in the effects of emotion, posture and place of measurement on blood pressure". In: *Social Science & Medicine* 31.11 (Jan. 1990), pp. 1213–1217. ISSN: 02779536. DOI: 10.1016/0277-9536(90)90126-D.

[157] Chris J. Boyatzis and Reenu Varghese. "Children's Emotional Associations with Colors". In: *The Journal of Genetic Psychology* 155.1 (Mar. 1994), pp. 77–85. ISSN: 0022-1325, 1940-0896. DOI: 10.1080/00221325.1994.9914760.

[158] M Lombard, Rd Reich, Me Grabe, Cc Bracken, and Tb Ditton. "Presence and television: The role of screen size". In: *Human Communication Research* 26.1 (Jan. 2000), pp. 75–98. ISSN: 0360-3989, 1468-2958. DOI: 10.1111/j.1468-2958.2000.tb00750.x.

[159] Emanuel Schmider, Matthias Ziegler, Erik Danay, Luzi Beyer, and Markus Bühner. "Is It Really Robust?: Reinvestigating the Robustness of ANOVA Against Violations of the Normal Distribution Assumption". In: *Methodology* 6.4 (2010), pp. 147–151. ISSN: 1614-1881, 1614-2241. DOI: 10.1027/1614-2241/a000016.

[160] Kwang-Kyu Seo, Sangwon Lee, Byung Do Chung, and Changsoon Park. "Users' Emotional Valence, Arousal, and Engagement Based on Perceived Usability and Aesthetics for Web Sites". In: *International Journal of Human-Computer Interaction* 31.1 (2015), pp. 72–87. ISSN: 1044-7318, 1532-7590. DOI: 10.1080/10447318.2014.959103.

[161] M Balconi and C Lucchiari. "Consciousness and arousal effects on emotional face processing as revealed by brain oscillations. A gamma band analysis". In: *International Journal of Psychophysiology* 67.1 (2008), pp. 41–46. ISSN: 01678760. DOI: 10.1016/j.ijpsycho.2007.10.002.

[162] Helder Bértolo, Teresa Paiva, Lara Pessoa, Tiago Mestre, Raquel Marques, and Rosa Santos. "Visual dream content, graphical representation and EEG alpha activity in congenitally blind subjects". In: *Cognitive Brain Research* 15.3 (2003), pp. 277–284. ISSN: 09266410. DOI: 10.1016/S0926-6410(02)00199-4.

[163] Alison Mattek. "Emotional communication in computer generated music: Experimenting with affective algorithms". In: *Proceedings of the 26th Annual Conference of the Society for Electro-Acoustic Music in the United States*. 2011.

[164] Jan-Niklas Antons. *Neural Correlates of Quality Perception for Complex Speech Signals*. 1st ed. 2015. T-Labs Series in Telecommunication Services. Cham: Springer International Publishing : Imprint: Springer, 2015. ISBN: 978-3-319-15521-0. DOI: 10.1007/978-3-319-15521-0.

[165] German Gomez-Herrero, Wim De Clercq, Haroon Anwar, Olga Kara, Karen Egiazarian, Sabine Van Huffel, and Wim Van Paesschen. "Automatic Removal of Ocular Artifacts in the EEG without an EOG Reference Channel". In: *Proceedings of the 7th Nordic Signal Processing Symposium - NORSIG 2006*. Rejkjavik: IEEE, 2006, pp. 130–133. ISBN: 978-1-4244-0412-4. DOI: 10.1109/NORSIG.2006.275210.

[166] Margaret M Bradley, Maurizio Codispoti, Bruce N Cuthbert, and Peter J Lang. "Emotion and motivation I: defensive and appetitive reactions in picture processing." In: *Emotion* 1.3 (2001), p. 276.

[167] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, PBC. Boston, MA, 2020. URL: http://www.rstudio.com/.

[168] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. "Fitting Linear Mixed-Effects Models Using lme4". In: *Journal of Statistical Software* 67.1 (2015), pp. 1–48. DOI: 10.18637/jss.v067.i01.

[169] Eva Maria Hammer, Sebastian Halder, Benjamin Blankertz, Claudia Sannelli, Thorsten Dickhaus, Sonja Kleih, Klaus-Robert Müller, and Andrea Kübler. "Psychological predictors of SMR-BCI performance". In: *Biological Psychology* 89.1 (2012), pp. 80–86. ISSN: 0301-0511. DOI: https://doi.org/10.1016/j.biopsycho.2011.09.006.

[170] Raoof Masoomi and Ali Khadem. "Enhancing LDA-based discrimination of left and right hand motor imagery: Outperforming the winner of BCI Competition II". In: *2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI)*. 2015, pp. 392–398. DOI: 10.1109/KBEI.2015.7436077.

[171] Andres Pinilla. *Contagio emocional: Incidencia de los estados emocionales sobre la imitación de expresiones faciales (Master's Thesis)*. Bogota, Colombia: Universidad Nacional de Colombia, 2017.

[172] Jonathan Peirce, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. "PsychoPy2: Experiments in behavior made easy". In: *Behavior Research Methods* 51.1 (Feb. 2019), pp. 195–203. ISSN: 1554-3528. DOI: 10.3758/s13428-018-01193-y.

[173] Christian Kruger, Tanja Kojic, Luis Meier, Sebastian Moller, and Jan-Niklas Voigt-Antons. "Development and Validation of Pictographic Scales for Rapid Assessment of Affective States in Virtual Reality". In: *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. Athlone, Ireland: IEEE, 2020, pp. 1–6. ISBN: 978-1-72815-965-2. DOI: 10.1109/QoMEX48832.2020.9123100.

[174] Arnaud Delorme and Scott Makeig. "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis". In: *Journal of Neuroscience Methods* 134.1 (Mar. 2004), pp. 9–21. ISSN: 01650270. DOI: 10.1016/j.jneumeth.2003.10.009.

[175] Alain Rakotomamonjy and Vincent Guigue. "BCI Competition III: Dataset II- Ensemble of SVMs for BCI P300 Speller". In: *IEEE Transactions on Biomedical Engineering* 55.3 (2008), pp. 1147–1154. DOI: 10.1109/TBME.2008.915728.

[176] Anne-Marie Brouwer, Thorsten O. Zander, Jan B. F. van Erp, Johannes E. Korteling, and Adelbert W. Bronkhorst. "Using neurophysiological signals that reflect cognitive or affective state: six recommendations to avoid common pitfalls". In: *Frontiers in Neuroscience* 9 (2015), p. 136. ISSN: 1662-453X. DOI: 10.3389/fnins.2015.00136. URL: https://www.frontiersin.org/article/10.3389/fnins.2015.00136.

[177] Klaus Gramann, Joseph T. Gwin, Daniel P. Ferris, Kelvin Oie, Tzyy-Ping Jung, Chin-Teng Lin, Lun-De Liao, and Scott Makeig. "Cognition in action: imaging brain/body dynamics in mobile humans". In: 22.6 (2011), pp. 593–608. DOI: doi:10.1515/RNS.2011.047.

[178] Massimiliano Palmiero and Laura Piccardi. "The Role of Emotional Landmarks on Topographical Memory". In: *Frontiers in Psychology* 8 (2017), p. 763. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2017.00763.

[179] Alejandro Ojeda, Nima Bigdely-Shamlo, and Scott Makeig. "MoBILAB: an open source toolbox for analysis and visualization of mobile brain/body imaging data". In: *Frontiers in Human Neuroscience* 8 (2014), p. 121. ISSN: 1662-5161. DOI: 10.3389/fnhum.2014.00121.

[180] Marius Klug, Lukas Gehrke, Hohlefeld Friederike U., and Klaus Gramann. *The BeMoBIL Pipeline - Facilitating Mobile Brain/Body Imaging (MoBI) Data Analysis in MATLAB*. Poster.

[181] Robert Soussignan. "Duchenne smile, emotional experience, and autonomic reactivity: A test of the facial feedback hypothesis." In: *Emotion* 2.1 (2002), pp. 52–74. ISSN: 1931-1516, 1528-3542. DOI: 10.1037/1528-3542.2.1.52.

[182] Guillaume Dezecache, Laurence Conty, Michele Chadwick, Leonor Philip, Robert Soussignan, Dan Sperber, and Julie Grèzes. "Evidence for Unintentional Emotional Contagion Beyond Dyads". In: *PLoS ONE* 8.6 (June 2013). Ed. by Manos Tsakiris, e67371. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0067371.

[183] Benjamin Blankertz, Guido Dornhege, Matthias Krauledat, Klaus-Robert Müller, and Gabriel Curio. "The non-invasive Berlin Brain–Computer Interface: Fast acquisition of effective performance in untrained subjects". In: *NeuroImage* 37.2 (2007), pp. 539–550. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2007.01.051.

[184] Irene Winkler, Mark Jäger, Vojkan Mihajlovic, and Tsvetomira Tsoneva. "Frontal EEG asymmetry based classification of emotional valence using common spatial patterns". In: *World Academy of Science, Engineering and Technology* 45 (2010), pp. 373–378.

[185] Tammy English, Oliver P. John, Sanjay Srivastava, and James J. Gross. "Emotion regulation and peer-rated social functioning: A 4-year longitudinal study". In: *Journal of Research in Personality* 46.6 (Dec. 2012), pp. 780–784. ISSN: 00926566. DOI: 10.1016/j.jrp.2012.09.006.

[186] Chadwick Boulay, Tristan Stenner, David Medine, Christian Kothe, Robert Guggenberger, Arnaud Delorme, Robert Oostenveld, Tim Mullen, Johnell, and Clement Lee. *sccn/labstreaminglayer*. Sept. 2021. URL: https://github.com/sccn/labstreaminglayer (visited on 09/14/2021).

[187] Yisi Liu, Olga Sourina, and Minh Khoa Nguyen. "Real-Time EEG-Based Emotion Recognition and Its Applications". In: *Transactions on Computational Science XII: Special Issue on Cyberworlds*. Ed. by Marina L. Gavrilova, C. J. Kenneth Tan, Alexei Sourin, and Olga Sourina. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 256–277. ISBN: 978-3-642-22336-5. DOI: 10.1007/978-3-642-22336-5_13.

[188] Yann Renard, Fabien Lotte, Guillaume Gibert, Marco Congedo, Emmanuel Maby, Vincent Delannoy, Olivier Bertrand, and Anatole Lécuyer. "OpenViBE: An Open-Source Software Platform to Design, Test, and Use Brain–Computer Interfaces in Real and Virtual Environments". In: *Presence: Teleoperators and Virtual Environments* 19.1 (Feb. 2010), pp. 35–53. ISSN: 1054-7460. DOI: 10.1162/pres.19.1.35.

[189] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A. Engemann, Daniel Strohmeier, Christian Brodbeck, Lauri Parkkonen, and Matti S. Hämäläinen. "MNE software for processing MEG and EEG data". In: *NeuroImage* 86 (2014), pp. 446–460. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2013.10.027.

[190] Mikel Val-Calvo, José R. Álvarez-Sánchez, Jose M. Ferrández-Vicente, and Eduardo Fernández. "Optimization of Real-Time EEG Artifact Removal and Emotion Estimation for Human-Robot Interaction Applications". In: *Frontiers in Computational Neuroscience* 13 (Nov. 2019), p. 80. ISSN: 1662-5188. DOI: 10.3389/fncom.2019.00080.

[191] A. D. I. Kramer, J. E. Guillory, and J. T. Hancock. "Experimental evidence of massive-scale emotional contagion through social networks". In: *Proceedings of the National Academy of Sciences* 111.24 (June 2014), pp. 8788–8790. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1320040111.

[192] Elaine Hatfield, John T. Cacioppo, and Richard L. Rapson. "Emotional Contagion". In: *Current Directions in Psychological Science* 2.3 (June 1993), pp. 96–100. ISSN: 0963-7214, 1467-8721. DOI: 10.1111/1467-8721.ep10770953.

[193] Guy Zyskind, Oz Nathan, et al. "Decentralizing privacy: Using blockchain to protect personal data". In: *2015 IEEE Security and Privacy Workshops*. IEEE. 2015, pp. 180–184.

[194] Wellington Fernandes Silvano and Roderval Marcelino. "Iota Tangle: A cryptocurrency to communicate Internet-of-Things data". In: *Future Generation Computer Systems* 112 (2020), pp. 307–319.

[195] Tortora Lucas. *iotaledger/stronghold.rs*. original-date: 2020-06-04T14:17:16Z. Aug. 2021. URL: https://github.com/iotaledger/stronghold.rs (visited on 08/17/2021).

[196] *Shapely*. original-date: 2011-12-31T19:43:11Z. Aug. 2021. URL: https://github.com/Toblerity/Shapely (visited on 08/18/2021).

A

Appendix A: Dimensional reduction

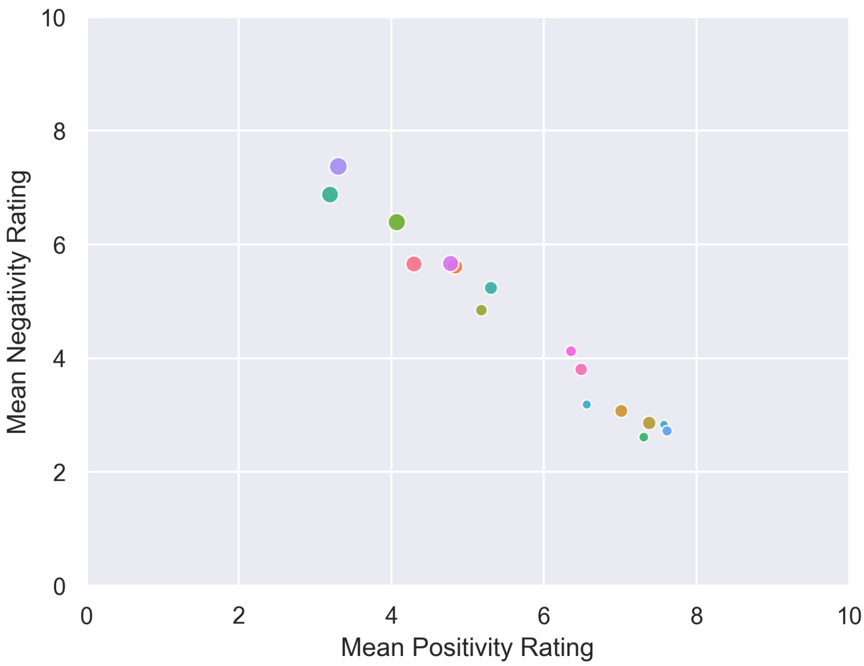
The DEAP dataset [23] is based on a two-dimensional model of affect: the Circumplex Model of Affect (CMA) [65]. In contrast, the pilot experiment described in section 4.2 was based on a three-dimensional model of affect: the Evaluative Space Model (ESM) [66]. Comparing the results of the DEAP dataset [23] with the results collected during the pilot experiment required mapping the results obtained in the three-dimensional model into a two-dimensional model. Therefore, a dimensional reduction method is proposed. The proposed method for dimensional reduction is based on two assumptions:

First assumption: The valence dimension of the CMA [65] is a bipolar dimension. Therefore, high values are associated with positive affective states and low values are associated with negative affective states. Consequently, the negativity and positivity dimensions of the ESM [66] are conceptually equivalent to the two poles of the valence dimension of the CMA [65].

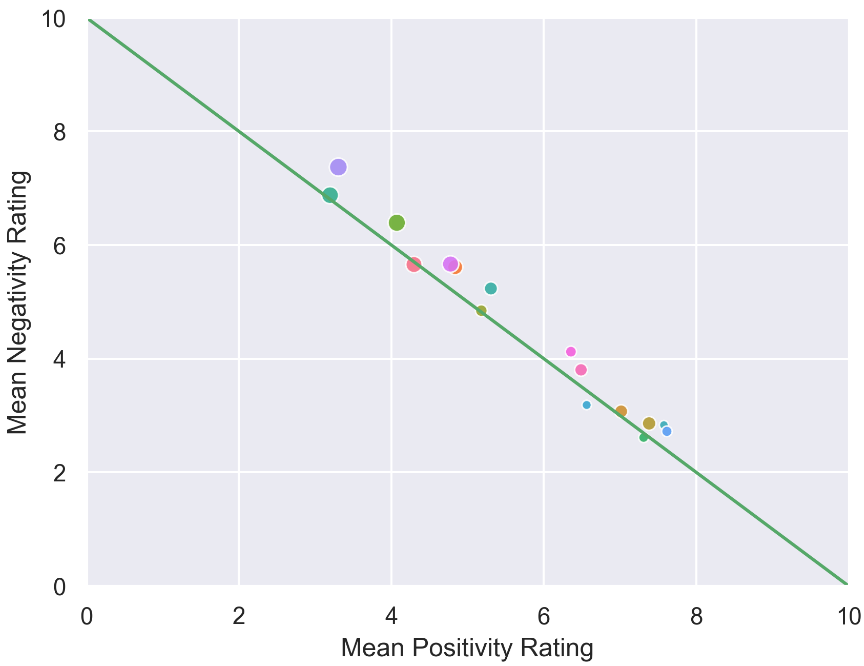
Second assumption: The net predisposition dimension of the ESM [66] is conceptually similar to the arousal dimension of the CMA [65]. The predisposition of an organism to approach or withdraw an stimuli is associated with the activation of its sympathetic system, hence associated with its level of arousal.

Based on those two assumptions, the proposed method consists of collapsing the negativity and positivity dimension of the ESM [66] into a single dimension. The net predisposition dimension of the ESM [66] is equivalent to the arousal dimension of the CMA [65]. The procedure is based on analytical geometry and consists of 5 steps:

Step 1. Plot the mean negativity and positivity ratings in a two dimensional space. The x-axis corresponds to the positivity ratings, while the y-axis corresponds to the negativity ratings.



Step 2. Trace a line that represents the valence dimension. High valence is conceptually equivalent to high positivity, while low valence is conceptually equivalent to high negativity. Therefore, valence can be represented with a diagonal line that connects the highest possible rating in the negativity dimension (upper left corner of the graph) with the highest possible rating in the positivity dimension (lower right corner of the graph).



The Equation of the Line is

$$y = x + b$$

where $\boldsymbol{x}$ is a value in the x-axis and $\boldsymbol{b}$ is the intercept (i.e., the value of $\boldsymbol{y}$ when $\boldsymbol{x} = 0$).

The intercept of the diagonal line that represents valence is 10. Therefore, the equation of the diagonal line that represents valence in the previous graph is

$$y = x - 10$$

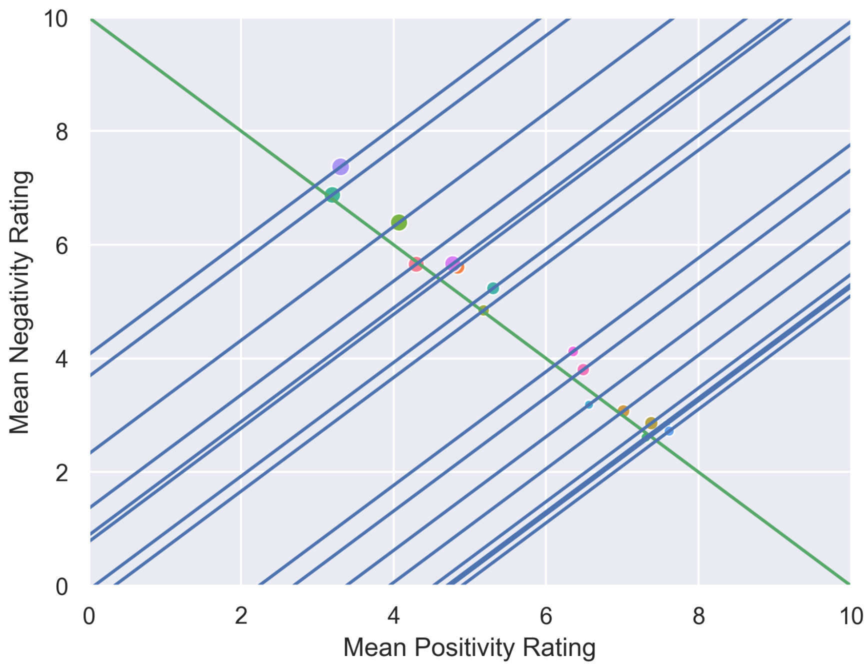
Step 3. Trace one line for each of the dots of the graph (each dot is a video). Each of those lines should be perpendicular to the diagonal line. The objective is to project each dot to the diagonal line.

The equation of a line that is perpendicular to the valence line is

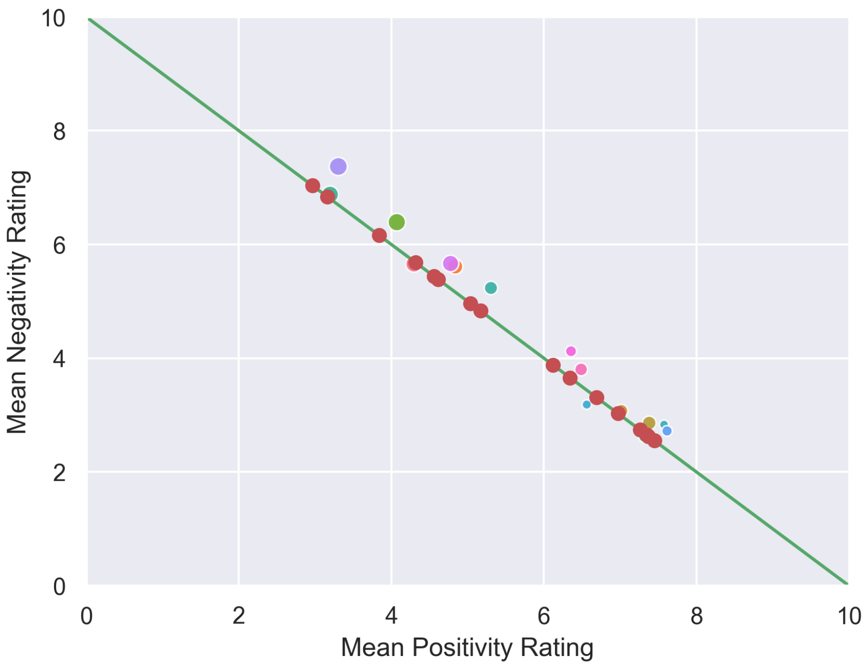
$$y = x - b$$

Therefore, any line traced with the previous equation should be perpendicular to the diagonal valence line traced in step 2.

Additionally, the $\boldsymbol{x}$ and $\boldsymbol{y}$ coordinates for all the dots in the graph can be inferred from participants' ratings. The ratings in the positivity scale are $\boldsymbol{x}$ coordinates, while the ratings in the negativity scale are the $\boldsymbol{y}$ coordinates. Therefore, it is possible to infer the intercept for each dot isolating $\boldsymbol{b}$. Once all the intercepts have been calculated, it is possible to trace the perpendicular line for each video.



Step 4. Find the coordinate where each perpendicular line intercepts the diagonal valence line. The coordinates of this intersection can be found using the *intersection()* function of the Shapely Python library [196].

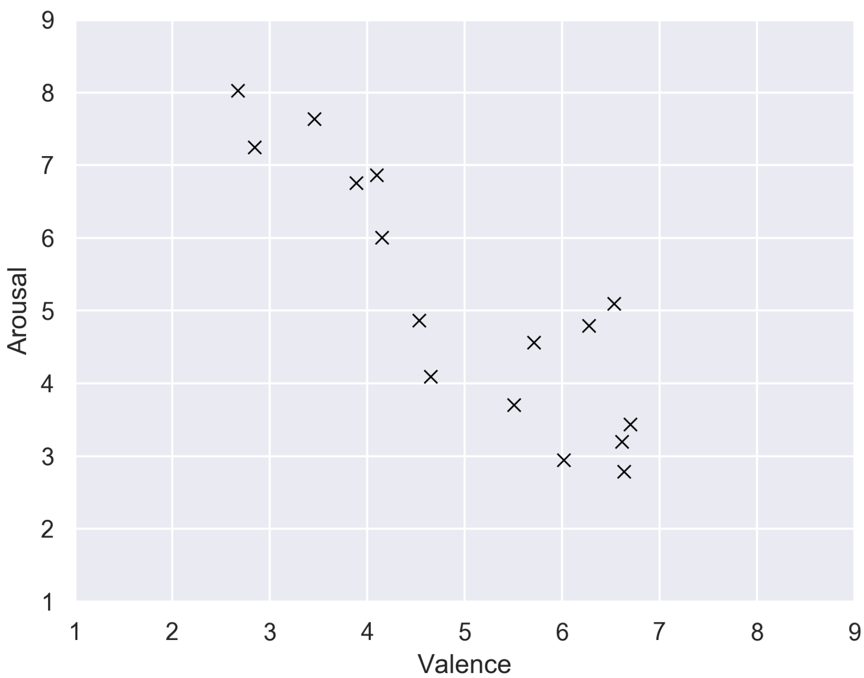


Step 5. The Pythagorean theorem can be used to calculate the distance between the intercept of the diagonal valence line (upper left side of the graph). Therefore, valence can be calculated with the equation

$$V = \sqrt{x^2 + y^2}$$

Where $\boldsymbol{x}$ and $\boldsymbol{y}$ are the values of the projected dots in the x-axis and y-axis, respectively (see step 4).

Step 6. Change range of the obtained values. The values obtained in the step 5 are in a different range than the valence ratings of the DEAP dataset [23]. Therefore, the range of the values obtained in the step 5 are adjusted for comparison purposes. A similar process is conducted for the Net Predisposition ratings. The range transformation is conducted using a linear conversion. The rescaled values are plotted in a valence-arousal space.



The ratings of the videos in the pilot experiment (see section 4.3) are compared with the ratings of the same videos in the DEAP dataset [23] (see section 4.1 and 4.2). Overall, when the dimensional reduction is conducted, the ratings of the videos in the pilot experiment (see section 4.3) are similar to the ratings of the DEAP dataset [23].