2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
17-20 October, 2021. Melbourne, Australia

Characterization of Affective States in Virtual Reality Environments using EEG

Meghan Kumar[1], Connor Delaney[1], Pedram Zanganeh Soroush[1], Yusuke Yamani[2], Dean J. Krusienski[1]

2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC) | 978-1-6654-4207-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/SMC52423.2021.9658916

***Abstract*— Recent interest in virtual reality (VR) headsets has motivated research efforts to increase the user's sense of immersion via feedback of physiological measures. This work presents the use of electroencephalographic (EEG) measurements during observation of immersive VR videos to estimate the user's affective state. A pilot was conducted on 10 participants. Participants passively viewed a series of one-minute immersive VR video clips and subjectively rated the level of valence, arousal, liking, and dominance. Correlates between EEG spectral bands and the subjective ratings were analyzed to identify statistically significant frequencies and electrode locations across participants.**

I. Introduction

The integration of user affective state feedback is critical to develop immersive closed-loop virtual reality (VR) experiences and improve human machine interaction (HMI) [1]. In recent studies, electroencephalogram (EEG) sensors have been integrated into VR headsets to measure human factors such as cognitive and affective state [2]. Room-scale wireless VR headsets are a practical tool to elicit affective state as they allow users to experience a realistic environment with naturalistic movements, while providing an unobtrusive framework for mounting EEG sensors [2], [3]. Additionally, the VR environmental parameters can be systematically varied to modulate cognitive and affective states [4].

Human emotion and affective states can be characterized using discrete and continuous models [5], [6]. A commonly used continuous model to evaluate a perceived affective state is the Russel Complex. The Russel Complex can be mapped into a two-dimensional space consisting of axes representing affective states, which are assigned numerical ratings. The ratings range from one to nine and the affective states are categorized into the dimensions of valence and arousal, with the extension of dimensions representing liking and dominance [7]. In this study, this extended Russel Complex was integrated into a Self-Assessment Manikin (SAM) to record the participant's affective state. SAM utilizes a nonverbal pictorial assessment technique based on the dimensions of the Russel complex [5], [8].

Traditional EEG analysis utilizes power spectral bands including the θ-band (4-7 Hz), α-band (8-13 Hz), β-band (14-29 Hz), and γ-band (30-47 Hz). These bands can be correlated with user subjective ratings across the frontal, central, parietal, and occipital locations. Koelstra et al. utilized music videos to elicit human emotional states and observed that an increase in valence was correlated to an increase in power for low frequency bands [7]. It was observed that an increase in the β-band power within right temporal regions was associated with positive emotional self-induction. In a study regarding music video excerpts, it was observed that higher frequency bands, such as the γ-band, were more prominent when subjects were listening to unfamiliar songs [9]. Sarno et al. have observed that high-frequency bands such as alpha, beta, and gamma are more effective for classifying emotions in both valence and arousal dimensions [10].

The present study aims to further the physiological characterization of human affective states in VR using EEG. A pilot study was performed on 10 participants using 13 publicly available 360° VR videos [11]. The EEG signals were measured while the participants viewed a series of VR videos, each followed by a prompt to complete a SAM. Correlates between the participant's affective state ratings and EEG frequency bands were computed. EEG electrodes found to be statistically significant in correlating with affective state were identified for future use to support the development of closed-loop VR systems.

II. Methodology

A. Participants and Experimental Setup

Ten healthy individuals (ages 19-34, mean 23.8) were recruited to participate in the experiment, which was approved by the Institutional Review Board of Virginia Commonwealth University. Each participant completed a screening process consisting of an informed consent form, demographic information form, and Motion Sickness Susceptibility Questionnaire (MSSQ) short form [12]. All participants satisfied the criteria by scoring a minimum of 19 on the MSSQ. Participants were provided guidelines to complete the SAM in terms of valence, arousal, dominance and liking prior to beginning experimental task [8].

The HTC VIVE hardware system consists of a motion-tracked headset display, and two "lighthouse" base stations that can provide 6 Degree of Freedom (6DOF) tracking. The VIVE wireless adapter was used in conjunction with the wireless EEG headset such that the participant was untethered and free to rotate to view the 360° videos. Following the participant screening, a visual calibration was performed using the HTC VIVE to correct for lens distance. The wireless 32-channel EEG cap was then placed on the participant's head and the EEG electrodes were filled with electrolyte gel. The electrode cap was covered with a

This work was supported by NSF 1944389.

[1]Meghan Kumar, Connor Delaney, Pedram Zanganeh Soroush, and Dean Krusienski are with the Biomedical Engineering Department, Virginia Commonwealth University, Richmond, VA, USA. kumarm2@vcu.edu

[2]Yuske Yamani is with the Psychology Department, Old Dominion University, Norfolk, VA, USA.

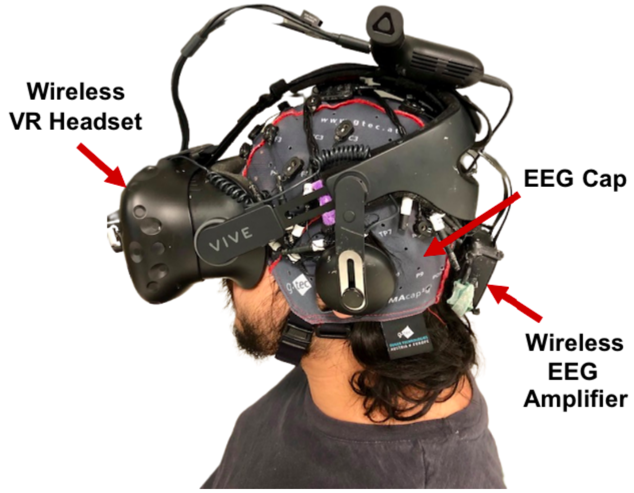

Fig. 1: Placement of wireless EEG cap and VR headset on participant.

protective plastic hair cap to protect the VR headset from the gel. The VR headset was placed over the EEG cap and the headset was tightened to comfortably fit the participant, as shown in Figure 1. Participants were then positioned approximately 1 meter from the recording computer in a seated position, in a swivel office chair.

B. Experimental Task

Stimuli are a series of 360° videos viewed through the VR headset to induce various affective reactions. The videos were obtained from a public database which were used in a previous human emotion study [11]. The videos from the database were segmented into three categories: High-Valence-Low-Arousal (HVLA), High-Valence-High-Arousal (HVHA), and Low-Valence-Low-Arousal (LVLA). Each category consisted of 4 one-minute videos for a total of 12 video segments, and one neutral video to collect baseline EEG data at the beginning of the experiment. At the start of the experiment, the participant viewed the neutral video and completed the SAM. The remaining 12 videos were then shown in a randomized order for each participant. For a particular trial, the participant's task was to passively view a video, with the ability to rotate the chair and the headset to view the 360° environment from different perspectives. Immediately following the video, participants were prompted to complete the SAM using a scale of one to nine for valence, arousal, dominance, and liking. The icons displayed for the SAM are shown in Figure 2 [8]. Each row represents a different affective state, and each column represents a different rating. The participant selects the ratings by directing his or her head toward the desired manikin and maintaining the cursor over the manikin for a 5-second dwell time, as indicated by a progress bar.

A twenty-second interval between trials was used to collect baseline EEG data. During this interval, participants were asked to keep their eyes open and remain still. The total duration of the experiment was kept to 25 min to reduce the risk of simulator sickness.

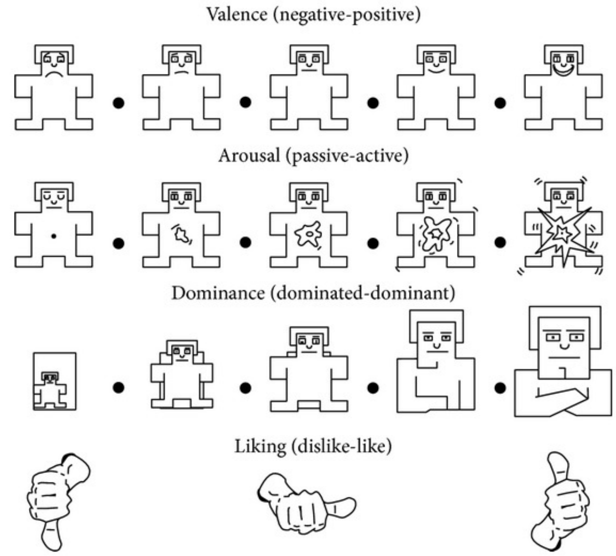

Fig. 2: Self-assessment manikin (SAM) for affective states. [15]

C. Data Collection

EEG data were collected using a 32-channel wireless biosignal amplifier (g.Nautilus, Guger Technologies) grounded to location AF3, referenced to the right earlobe, and sampled at 250 Hz with a bandpass filter from 0.1-100 Hz and a notch filter between 58-62 Hz. The electrode positions are based on the International 10-20 system as shown in Figure 3.

Communication between the VR environment (developed in Unity [13]) and the EEG recording was performed via Lab Streaming Layer [14] and recorded using the Lab Recorder application.

D. Data Analysis

The EEG data were parsed into 12 one-minute trials per participant. To correct for unrelated stimulus variation in power, the baseline EEG signal from the 20 second inter-trial interval was used. The EEG spectral power between 2 Hz and 47 Hz was computed of the first 30 seconds of each trial and baseline was computed using Welch's method on non-overlapping 250-sample segments with a 250-point Fast Fourier Transform (FFT), yielding 1-Hz spectral bins. The baseline power was subtracted from the trial power, yielding the change of power relative to pre-stimulus period. The net changes of power spectral density were then averaged over four frequency bands: θ (4-7 Hz), α (8-13 Hz), β (14-29 Hz), and γ (30-47 Hz).

For the correlation analysis, Spearman coefficients between the power changes and the subjective ratings were computed for each channel, along with the p-values for positive and negative correlation tests. This was done for each participant separately and, assuming independence,

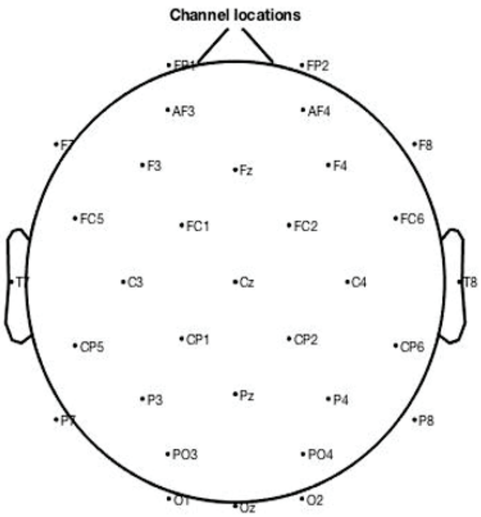

Fig. 3: The EEG cap channel locations electrode locations using the International 10-20 system.

Fig. 4: Boxplot of affective state ratings for HVLA videos.

the 32-resulting p-values per correlation direction (positive/negative), frequency band, and electrode were combined to one p-value using Fisher's method [7], [10].

III. Results

A. Affective State Ratings

Figures 4-6 show the distribution of ratings for the three categories of videos observed: HVLA, HVHA, LVLA, respectively. For the HVLA video ratings (Figure 4) valence and liking had a median of 8, larger than the median of 6 for arousal and 4 for dominance. Videos in the HVHA video category (Figure 5) had a median of 8 for valence, arousal, and liking, and 6 for dominance. Videos categorized as LVLA (Figure 6) had comparatively low medians ranging from 3-4.5 for valence, arousal, and liking. The median value for dominance was 7 for LVLA, suggesting the videos are more influential than the other categories. While the median values correspond to the affective state categorizations as determined by [11], for each categorization it is noted that the range of ratings is consistently large for all affective states.

Fig. 5: Boxplot of affective state ratings for HVHA videos.

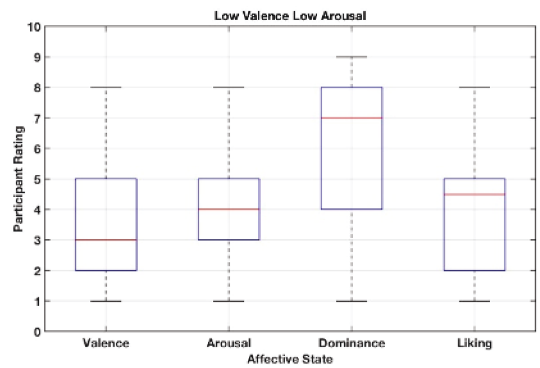

Fig. 6: Boxplot of affective state ratings for LVLA videos.

TABLE I: The means of the subject-wise intercorrelations between the scales of valence, arousal, liking, dominance, for all 10 participants. Significant correlations ($p \leq 0.05$) using Fisher's method are indicated by *.

	Valence	Arousal	Dominance	Liking
Valence	1	0.46*	0.12*	0.83*
Arousal	-	1	0.58*	0.50*
Dominance	-	-	1	0.31*
Liking	-	-	-	1

The mean intercorrelation of the different scales over the 10 participants (see Table I) was explored to indicate possible confounds or unwanted effects of fatigue from certain video categories. All intercorrelations were found to be statistically significant, with a large positive correlation between liking and valence (ρ = 0.83). Correlations ranging from ρ = 0.46 to ρ = 0.58 were observed between valence/arousal, dominance/arousal, and between arousal/liking. The lowest intercorrelation was observed between valence and dominance (ρ = 0.12).

B. Bandpower Correlation with Affective State

Figure 7 shows topographies of the average Spearman correlations with significantly correlated electrodes highlighted in grey ($p \leq 0.05$). The comprehensive list of significant electrodes is provided in Table II. A total of 3 electrodes yielded $p \leq 0.01$. Significant electrodes across all bands largely consisted of electrodes located in the frontal region.

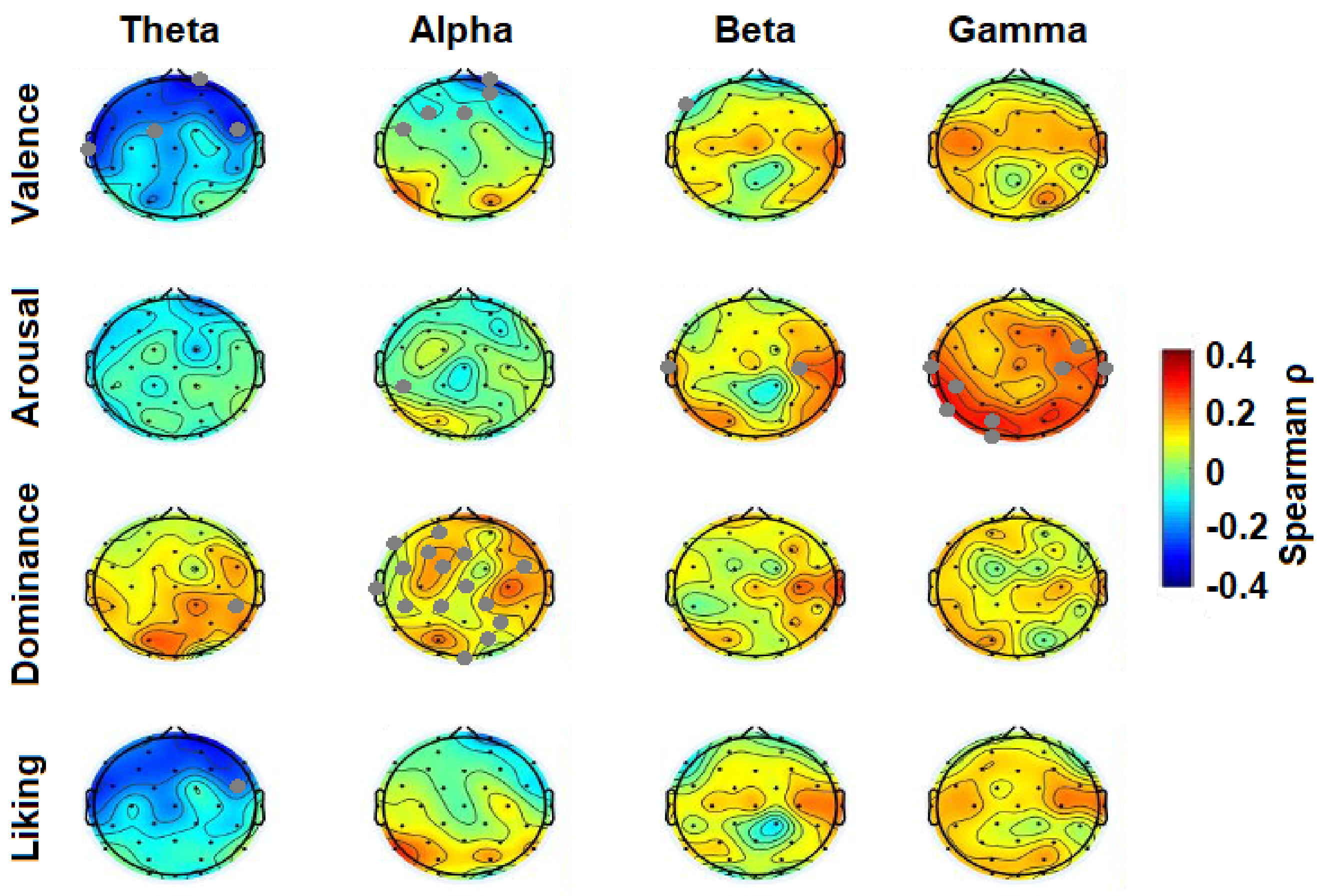

Fig. 7: The mean Spearman correlations over 10 participants between valence, arousal, dominance, and liking with the power in the frequency bands of θ (4-7 Hz), α (8-13 Hz), β (14-29 Hz), and γ (30-47 Hz), respectively. The grey highlighted electrodes correlate significantly ($p \leq 0.05$) with ratings.

TABLE II: Electrodes exhibiting significant correlation between bandpower and scale ($p \leq 0.05$, ***$p \leq$0.01**).

	Valence	Arousal	Dominance	Liking
Theta	FC1, T7, **FC6***, FP2	-	CP6	FC6
Alpha	F3, FC5, AF4, FP2, Fz	CP5	**T7***, AF3, F7, F3, FC1, FC5, CP1, CP5, Oz, PO4, P4, CP2, FC6, Fz, Cz	-
Beta	F7	T7,C4	-	-
Gamma	-	**T7***, CP5, P7, PO3, O1, C4, T8, FC6	-	-

T7 is a common electrode within all bands.

There is a significant increase in negative correlation within the frontal region for the theta band. There is a positive correlation for theta dominance towards the occipital region with a decrease towards the frontal region. Alpha valence and liking follow similar trends with negative correlation in the right frontal region and have an increase in positive correlation in left parietal region. For beta valence, arousal, and liking there is central negative correlation, with an increase in positive correlation in the right temporal/parietal region. The gamma band exhibits a positive correlation in the temporal/parietal regions for all rating categories. Correlates were found towards the frontal region in valence ratings, and near temporal region for arousal ratings. There are 15 significant electrodes for the alpha band dominance rating, with none for the beta and gamma bands.

IV. Discussion

The reported affect scores generally aligned with the affective state categorizations HVHA, LVLA, and HVLA. The HVHA category exhibited the highest median rating across valence, arousal, and liking. This supports the notion that more pleasing videos are rated above 5 for each state, indicating the participant felt positive emotion while watching these videos. Videos in the LVLA category had a median lower than 5 for valence, arousal, and liking indicating videos intended to elicit negative emotion were rated as unpleasant and were generally disliked. Dominance rating signifies the feeling of insignificant versus influential. For the category of LVLA, the higher dominance median can be correlated to the impactful video excerpts shown to the participant. Videos meant to elicit a calm/relaxed feeling were presented in the HVLA category. The low dominance median rating and lower arousal rating can be correlated to a calm state within HVLA. The ratings for the valence and liking for HVLA suggest these videos elicited a positive feeling.

The intercorrelations in Table I indicate that there are significant relationships between most pairs of affect scores. In terms of valence and arousal, this is likely a function of the prescribed categorizations, with the HVHA and LVLA creating inherent correlation. The comparatively large correlation between valence and liking is unsurprising since positive/negative emotions (valence) are generally associated with liking/disliking, respectively [7]. As expected from the intercorrelation analysis, the topographies for valence and liking are very similar across frequency bands.

The two main areas of the brain that are generally correlated with emotion and affective state are the amygdala and the frontal lobe. The statistically significant electrodes indicated in Figure 7 and Table II show that frontal alpha is prominent for valence and alpha is more distributed for dominance. Studies have shown that the frontal region reflects greater emotional activation compared to other regions of the brain such as the temporal, parietal, and occipital regions [9]. Electrode T7 in the temporal region was statistically significant with $p \leq 0.01$ for dominance in the alpha band and arousal in the gamma band. The temporal region has previously been associated with positive emotional self-induction and external stimulation [7]. A positive correlation of arousal, particularly in the gamma band, emanating from the anterior temporal region have also been reported [11]. However, electromyographic (EMG) activity is also known to be prominent in the higher frequencies over anterior and temporal electrodes [7], particularly as a result of emotive facial expressions [16]. While EMG was not obvious during visual inspection of the signals, modulations due to more subtle muscle tension cannot be ruled out. Similarly, it is also possible that eye movement artifacts, which were not explicitly measured or analyzed in this study, may contribute to the correlations observed in the frontal locations.

This pilot study on 10 participants yielded promising, statistically-significant results in term of identifying EEG channels and frequency bands that correlate with affective states. Due to the subjective nature of the task and varying style and content of the available 360° videos, further analysis needs to be performed on the impact of the specific video stimuli (e.g., motion, audio, realism, optical flow, etc.) on EEG. Future work will include a larger number of participants and an investigation of the potential for online classification of affective state via EEG. Ultimately, it is believed that such physiological feedback will greatly enhance the VR user-experience.

References

[1] R. Cassani, M.-A. Moinnereau, L. Ivanescu, O. Rosanne, and T. Falk, "Neural interface instrumented virtual reality headsets: Toward next-generation immersive applications," *IEEE Systems, Man, and Cybernetics Magazine*, vol. 6, pp. 20–28, 07 2020.

[2] C. Tremmel, C. Herff, and D. J. Krusienski, "EEG movement artifact suppression in interactive virtual reality," *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019.

[3] J. Blascovich, J. Loomis, A. C. Beall, K. R. Swinth, C. L. Hoyt, and J. N. Bailenson, "Target article: Immersive virtual environment technology as a methodological tool for social psychology," *Psychological Inquiry*, vol. 13, no. 2, p. 103–124, 2002.

[4] C. Tremmel, C. Herff, T. Sato, K. Rechowicz, Y. Yamani, and D. J. Krusienski, "Estimating cognitive workload in an interactive virtual reality environment using EEG," *Frontiers in Human Neuroscience*, vol. 13, 2019.

[5] J. A. Russell, "A circumplex model of affect." *Journal of Personality and Social Psychology*, vol. 39, no. 6, p. 1161–1178, 1980.

[6] R. Plutchik, "The nature of emotions: Clinical implications," *Emotions and Psychopathology*, p. 1–20, 1988.

[7] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, p. 18–31, 2012.

[8] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25, no. 1, p. 49–59, 1994.

[9] R. Sarno, M. N. Munawar, and B. T. Nugraha, "Real-time electroencephalography-based emotion recognition system," *International Review on Computers and Software (IRECOS)*, vol. 11, no. 5, p. 456, 2016.

[10] N. A. Lazar, B. Luna, J. A. Sweeney, and W. F. Eddy, "Combining brains: A survey of methods for statistical pooling of information," *NeuroImage*, vol. 16, no. 2, p. 538–550, 2002.

[11] B. J. Li, J. N. Bailenson, A. Pines, W. J. Greenleaf, and L. M. Williams, "A public database of immersive vr videos with corresponding ratings of arousal, valence, and correlations between head movements and self report measures," *Frontiers in Psychology*, vol. 8, 2017.

[12] J. F. Golding, "Predicting individual differences in motion sickness susceptibility by questionnaire," *Personality and Individual Differences*, vol. 41, no. 2, p. 237–248, 2006.

[13] "Unity Real-time Development Platform," https://unity.com/.

[14] "Lab Streaming Layer," https://github.com/sccn/labstreaminglayer.

[15] J. Morris, "Observations: SAM: The self-assessment manikin an efficient cross-cultural measurement of emotional response 1," *Journal of Advertising Research*, 1995.

[16] J. Lou, Y. Wang, C. Nduka, M. Hamedi, I. Mavridou, and H. Yu, "Realistic facial expression reconstruction for vr hmd users," *IEEE Transactions on Multimedia*, vol. PP, pp. 1–1, 08 2019.