2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

An Affective Interaction System using Virtual Reality and Brain-Computer Interface

Zheng Yang Chin, Zhuo Zhang, Chuanchu Wang, Kai Keng Ang

Abstract— **Affective Computing is a multidisciplinary area of research that allows computers to perform human emotion recognition, with potential applications in areas such as healthcare, gaming and intuitive human computer interface design. Hence, this paper proposes an affective interaction system using dry EEG-based Brain-Computer Interface and Virtual Reality (BCI-VR). The proposed BCI-VR system integrates existing low-cost consumer devices such as an EEG headband with frontal and temporal dry electrodes for brain signal acquisition, and a low-cost VR headset that houses an Android handphone. The handphone executes an in-house developed software that connects wirelessly to the headband, processes the acquired EEG signals, and displays VR content to elicit emotional responses. The proposed BCI-VR system was used to collect EEG data from 13 subjects while they watched VR content that elicits positive or negative emotional responses. EEG bandpower features were extracted to train Linear Discriminant and Support Vector Machine classifiers. The classification performances of these classifiers on this dataset and the results of a public dataset (SEED-IV) are then evaluated. The results in classifying positive vs negative emotions in both datasets (~66% for 2-class) show promise that positive and negative emotions can be detected by the proposed low cost BCI-VR system, yielding nearly the same performance on the public dataset that used wet EEG electrodes. Hence the results show promise of the proposed BCI-VR system for real-time affective interaction applications in future.**

Index Terms— Emotion detection; EEG; Virtual reality (VR); Affective computing (AfC)

I. INTRODUCTION

As one of the most fundamental mental processes, emotions play an essential role in a person's interaction with the external world. Hence a goal of Affective Computing is to develop socially smart Human Computer Interface (HCI) systems that deciphers and responds to the emotional states of the user [1]. A person's external expressions can be measured through audio/visual signals, while inner emotional state are assessed with self-reports or via physiological signals [2] such as the non-invasive electroencephalogram (EEG) signals, since emotions originate from the cerebral cortex. Hence, identifying EEG correlates of emotions complements subjective self-reports by providing a window into the inner emotional state of the individual. Furthermore, EEG has a relatively higher temporal resolution compared to other brain signal acquisition methods such as functional near infrared imaging (fNIRS) or functional magnetic resonance imaging (fMRI). In recent years, lightweight dry EEG headsets are available to everyday consumers too, making it more accessible for researchers to study the identification of EEG correlates in emotion recognition studies.

Eliciting emotions in human emotion recognition studies can be categorized into *active* and *passive* modes based on the nature of the stimuli [3]. In active elicitation methods, individuals could be instructed to adopt certain behaviours or facial expressions that might evoke different affective states naturally. In passive emotional elicitation, stimuli which are designed to evoke different emotions are presented. Such stimuli could be standardized to ensure all individuals have the same viewing experience. They are usually presented as:

1) *Audio stimuli* – such as the IADS (International Affective Digitalised Sound System) [4] database , comprising more than 100 sounds categorized along the affective dimensions of valence, arousal and dominance.

2) *visual stimuli* – such as the IAPS (International Affective Picture System) database [5], which has been used as an elicitation source for emotion recognition research.

3) audio-visual stimuli – such as the SEED-IV (SJTU Emotion EEG IV Dataset)[6], which uses films to elicit various levels of valence and arousal. Films present dynamic visual and auditory stimuli that may bear more similarities to real life scenarios [7]. In this dataset, EEG data is recorded from a full scalp gel-based wet EEG device.

An important factor for such stimuli in emotion elicitation is the degree of immersion that individuals experience. Research has shown that self-reported intensity of emotion is significantly greater in immersive than in non-immersive environments, with the presentation of the same content [8]. An example of an immersive environment is virtual reality (VR), which isolates the user from external world interferences, and enhances the immersive experience to induce emotions through the simulation of real experiences. With advances in computer graphics technology, the display of immersive VR content on mobile phones is possible, thus making it more accessible to everyday consumers.

In [9], the authors explored recording wet EEG and electrocardiogram(ECG) signals while subjects viewed VR content using a Samsung Gear VR headset. Nine EEG

[1] Research supported by Agency for Science, Technology and Research, (A*STAR), Singapore. Z. Y. Chin, Z. Zhang, C. Wang, K. K. Ang are with Institute for Infocomm Research, A*STAR, 1 Fusionopolis Way #21-01 Connexis (South Tower) Singapore 138632 Email: {zychin, zzhang, ccwang, kkang}@i2r.a-star.edu.sg. K.K. Ang is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore kkang@ntu.edu.sg

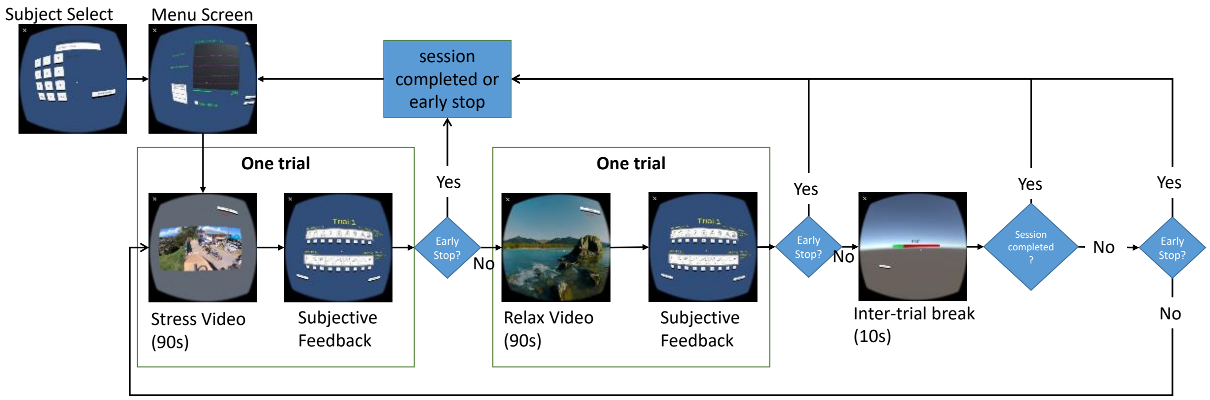

Figure 1 : Experiment protocol of the proposed BCI-VR system evaluation for emotion classification. In each trial, the subject watches a video. In the first trial, a negative emotion (stress) video lasting 90s is presented. The subject then provides subjective feedback. In the next trial, a positive emotion (relax) video lasting another 90s is presented, followed by another subjective feedback. At the end, a 10s break period is given before the next trial begins.

electrodes were placed in the frontal, central and parietal regions. Combining EEG and ECG features with their SVM classifier yielded around 70% to 71% test set accuracy to differentiate positive and negative valence, and positive and negative arousal. The results from [9] provide motivation in this paper to explore using commercially available dry EEG headsets for BCI-VR applications. Hence, this paper seeks to: investigate eliciting emotional responses using a proposed BCI-VR mobile device system, where the it integrates existing low cost consumer devices: an dry EEG headband and a VR headset that houses an Android handphone. To evaluate its feasibility, an in-house EEG data collection would be conducted where subjects used the BCI-VR system while viewing different VR content. EEG classification models are trained to discriminate two different emotions elicited. To benchmark the performance of the models and the proposed BCI-VR system, the EEG data from the SEED-IV public dataset will analysed.

II. Methodology

A. Proposed BCI-VR System

The proposed system is shown in Figure 2.The hardware comprises a Muse (https://choosemuse.com/) EEG headband with forehead (AF1, AF2) and temporal (TP9, TP10) dry electrodes and a Google Daydream VR headset that houses an Android handphone. The software deployed in Android device includes user interface (UI) and data communication components. The UI component is a customized Unity3D application that displays VR content and solicits subjective feedback from the subject at the end of each movie. The feedback is controlled purely by interacting with the screen content with head movement thus no keyboard or mouse is needed. The VR content displays a virtual cinema that shows normal video content or VR360 videos. VR360 videos, also known as 360-degree video or immersive videos, are video recordings where a view in every direction is recorded at the same time with an omnidirectional camera or a collection of cameras. During playback on normal flat display the viewer has control of the viewing direction like a panorama. The communication software is an extension of the Neurocomm Platform [10] for the Android platform. It communicates with EEG headset via Bluetooth to received EEG signals at 256 Hz sampling rate. The EEG signals and the associated events that occurred during the experiment are stored in the handphone for further analysis.

B. Emotion experiment protocol

The experimental procedures involving human subjects described in this paper were approved by the Institutional Review Board (2018-008). 13 subjects participated in the experiment. The experiment protocol elicits emotional responses as shown in Figure 1. There were 20 trials in total, where each trial presented a video; which was either eliciting a negative (stress) or positive (relax) emotion in the subject. These videos were curated from YouTube VR360 video clips. Each video lasted ~90s.

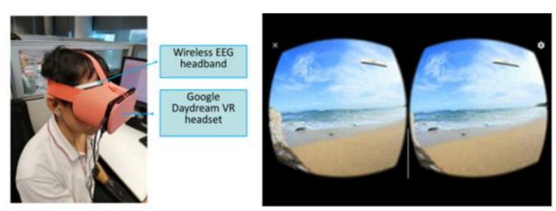

Figure 2: (left) Proposed BCI-VR system integrates a commercially available EEG headband and a VR headset which houses a handphone displaying VR content to the subject. (right) sample VR movie showing a beach scene

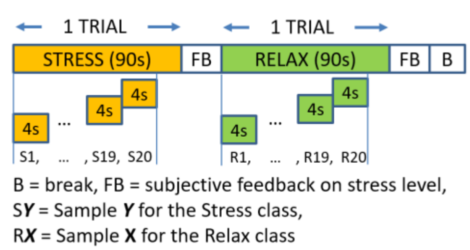

Figure 3: EEG segmented in each trial to form two classes for emotion classification; Positive (Relax) versus Negative (Stress) emotions. Non-overlapping 4s segments are extracted, resulting in 20 samples from each trial

C. EEG data processing

The continuous EEG data was visually inspected offline using MNE Python [11]. Noisy EEG electrodes were removed from two subjects for subsequent analysis. The EEG data was divided into two classes: positive and negative emotions elicited while viewing the stress videos and relax videos.

1) Filtering and data splitting

A highpass filter at 0.2 Hz and a notch filter of 50Hz were applied to remove baseline drifts and power line noise respectively. The raw EEG data was segmented into 20 trials of 90s each. The 20 trials were split into training (12), validation (4) and test (4) sets. The splitting is carried out at the trial level to avoid data leakage contamination among the three sets during hyperparameter tuning and classification.

2) Segmenting into 4-second windows for classification

Five bandpass filters (1-4Hz, 4-8Hz, 8-14Hz, 14-31Hz and 31-50Hz) were then applied onto each 90s trial. Each trial was split into 4s non-overlapping samples for classification as shown in Figure 3. This yielded 22 samples (22 x 4s = 88s) from each trial. For each subject, the training set, validation set and test set comprised 264, 88 and 88 samples.

D. Model Training

1) Log bandpower feature extraction and selection

EEG log bandpowers were computed from each sample, yielding 20 features (5 bandpass filters x 4 electrodes). Feature selection using Fisher Ratio (FR) is subsequently employed to select a set of most discriminative features, i.e. features with the highest Fisher Ratio (FR) for classification. FR is defined as the ratio of the variance of the between classes S_B to the variance of the within classes S_w, shown in (1) and (2).

$$\mathrm{FR} = S_B / S_W \tag{1}$$

$$S_B = \sum_{k=1}^{C} \frac{n_k}{n} (m_k - m)^2 , S_W = \sum_{k=1}^{C} \frac{n_k}{n} \sum_{j=1}^{n_k} \left(x_{kj} - m_k\right)^2 \tag{2}$$

where C is the total number of classes, n_k is the number of trials in class k, n is the total number of trials, m_k is the mean of the feature for class k, m is the overall mean of the feature, x_{kj} is the feature for the j-th trial in class k. The number of selected features is a hyperparameter that was tuned during the classifier model training.

2) Classification

Two classifiers were evaluated: Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM). For the SVM classifier, the hyperparameters included the C value, the Gamma value and kernel type (radial basis function or sigmoid). For LDA, the hyperparameter included the Shrinkage value. The models were first trained using data from the training set and hyperparameters were tuned using the validation set, where an exhaustive search (Grid Search) approach was employed. The best hyperparameters were selected based on the best validation set accuracy. Finally, the model is retrained with the best hyperparameters on the training and validation set, and a final evaluation result on the test set is presented.

E. Comparison of results with existing public dataset

Analysis with a publicly available dataset, SEED-IV [6] was also carried out to compare the usage of the same EEG processing methods with wet EEG electrodes. The SEED-IV dataset comprises data from 15 subjects, watching film clips which induces happy, sad, fear or neutral emotions. There were 72 trials per subject. EEG data was recorded from 62 wet electrodes using the Neuroscan EEG amplifier. The raw EEG data from FP1, FP2, T7 and T8 in the dataset was processed using the same methods as described earlier. These electrodes best match the electrode positions used in the MUSE headset. For each subject, the number of trials in the training set, validation set and test set were 36, 24 and 12 respectively.

III. Results and Discussion

This section presents analysis results for data collected using the proposed system (denoted as BCI-VR) and the SEED-IV dataset.

TABLE I: 2-CLASS CLASSIFICATION ACCURACIES ON TRAINING, VALIDATION AND TEST SETS FOR BOTH LDA AND SVM FOR BCI-VR DATASET

2-class	***LDA***			***SVM***		
Subject	training	valuation	testing	training	valuation	testing
1	0.686	0.773	0.500	0.788	0.761	0.455
2	0.678	0.739	0.739	0.693	0.693	0.693
3	0.678	0.576	0.491	0.570	0.606	0.436
4	0.678	0.705	0.716	0.826	0.807	0.705
5	0.735	0.716	0.807	0.777	0.773	0.761
6	0.731	0.727	0.545	0.648	0.750	0.727
7	0.693	0.614	0.614	1.000	0.682	0.614
8	0.670	0.807	0.784	0.678	0.830	0.807
9	0.754	0.739	0.773	0.723	0.727	0.795
10	0.773	0.716	0.716	0.769	0.705	0.716
11	0.701	0.625	0.636	0.917	0.670	0.727
12	0.735	0.705	0.511	0.830	0.818	0.511
13	0.886	0.568	0.602	0.852	0.614	0.580
mean	0.723	0.693	**0.649**	0.775	0.726	**0.656**
std	0.059	0.074	**0.114**	0.115	0.073	**0.125**

TABLE II: 4-CLASS CLASSIFICATION ACCURACIES ON TRAINING, VALIDATION AND TEST SETS FOR BOTH LDA AND SVM FOR SEED-IV PUBLIC DATASET

4-class	***LDA***			***SVM***		
Subject	training	valuation	testing	training	valuation	testing
1	0.465	0.415	0.326	0.652	0.424	0.300
2	0.592	0.569	0.562	0.709	0.600	0.572
3	0.469	0.448	0.233	0.677	0.519	0.349
4	0.610	0.507	0.538	0.957	0.510	0.531
5	0.558	0.476	0.372	0.629	0.510	0.331
6	0.422	0.446	0.346	0.635	0.514	0.395
7	0.558	0.559	0.497	0.547	0.541	0.444
8	0.590	0.505	0.538	0.702	0.554	0.608
9	0.649	0.526	0.369	0.745	0.541	0.431
10	0.552	0.468	0.444	0.623	0.549	0.446
11	0.485	0.442	0.341	0.890	0.443	0.469
12	0.438	0.380	0.372	0.856	0.402	0.379
13	0.532	0.416	0.413	0.586	0.420	0.400
14	0.576	0.606	0.433	0.731	0.641	0.505
15	0.711	0.453	0.590	0.853	0.534	0.615
mean	0.547	0.481	**0.425**	0.720	0.513	**0.452**
std	0.080	0.063	**0.102**	0.120	0.067	**0.098**

A. 2-class results from proposed BCI-VR

Table I shows the 2-class (negative emotion / positive emotion) classification accuracies on the training, validation and test set. The SVM classifier with optimized hyperparameters yielded higher classification accuracies than the LDA classifier on average over all subjects for all three sets. Though not directly comparable, the results(~66%) are slightly lower than the results in [9] (~70%). In [9], both EEG and ECG signals were recorded, where 9 wet EEG electrodes were placed over the frontal, central and parietal regions, with a pair of electrodes below the mastoid as reference, and a pair of ECG leads on the rib and collarbone.

B. Comparing results from BCI-VR and SEED-IV

Table II shows the 4-class (happy, sad, fear, neutral) classification accuracies on the training set, validation set and test set of the SEED-IV dataset. Similar to the results in Table I, the SVM classifier yielded higher accuracies across all three sets too.

For a more direct comparison between the results from the BCI-VR dataset and the SEED-IV dataset, the four classes from the SEED-IV dataset were combined to form two classes: negative emotion (sad, fear) and positive emotion (happy, neutral). The confusion matrices and classification accuracies for SVM on the test sets are shown in Table III. The accuracies for the majority class classifier (naïve) are also presented. The 2-class SVM classification accuracy results on the test sets for both BCI-VR and SEED-IV are similar and is higher than the naïve accuracies. The results suggest that existing EEG processing algorithms could classify positive and negative emotions elicited by proposed BCI-VR setup with the dry EEG headset, around the same performance by wet EEG electrodes. The temporal region of the brain is responsible for emotions while the frontal region is associated with cognitive processing [12] thus the TP9 and TP10 electrodes in the dry EEG headset could have contributed to the classification performance. In [6], the authors also identified the temporal electrodes in lieu of the 62-channel Neuroscan EEG amplifier for use in their device as well.

Nonetheless, there are limitations in the current study. First, the subjective feedback was not accounted for yet. Second, the current prediction is based on binary classification. Third, the EEG signals could be easily contaminated artefacts such as EOG and EMG signals. Hence, future work could include exploring how to integrate he subjective feedback, a continuous score for emotion categorization and identifying how the EEG could be contaminated.

TABLE III: 2-CLASS SVM CONFUSION MATRICES ON TEST SET FOR BCI-VR AND SEED-IV. THE CLASSIFICATION ACCURACIES IN BOTH DATASETS ARE SIMILAR AT 0.65~0.66. THE NAÏVE ACCURACIES ARE COMPUTED USING A MAJORITY CLASS CLASSIFIER

dataset		*Predicted (%)*		*Accuracy*	
		Positive	*Negative*	***SVM***	**Naïve**
BCI-VR	*Positive*	0.59	0.41	0.65	0.51
	Negative	0.29	0.71		
SEED-IV	*Positive*	0.65	0.35	0.66	0.54
	Negative	0.33	0.67		

IV. CONCLUSION

This paper proposed and investigated the performance of a low-cost EEG-based BCI-VR mobile device system which integrates an existing dry EEG acquisition device and a VR headset that presents audiovisual stimuli using an Android handphone. Compared to an off-the-shelf integrated EEG and VR system, the design of the proposed BCI-VR system enables a cost-effective and convenient experiment to elicit emotions for Affective Computing studies. To evaluate the system, an in-house data collection was carried out on 13 subjects, who viewed positive and negative VR content while 4-channel EEG data was collected. EEG-based classification using EEG log bandpower features and SVM yielded a validation and test accuracy of about 73% and 66% for 2-classes (positive versus negative) respectively. The test set are similar to the 2-class test set results obtained on the SEED-IV public dataset. The similarity in classification performance suggest that the EEG processing algorithms could classify positive and negative emotions elicited by the proposed BCI-VR setup with the dry EEG headset, at around the same performance on the public dataset which uses wet EEG electrodes. Hence, the results provide motivation to extend the functionality of the proposed BCI-VR system to include real-time emotion recognition in future.

REFERENCES

[1] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 1997.

[2] H. Gunes, B. Schuller, M. Pantic, and R. Cowie, “Emotion representation, analysis and synthesis in continuous space: A survey,” in *2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, Mar. 2011, pp. 827–834, doi: 10.1109/FG.2011.5771357.

[3] J. Kory and S. D’Mello, “Affect Elicitation for Affective Computing,” in *The Oxford Handbook of Affective Computing*, Oxford University Press, 2014, pp. 371–383.

[4] M. Nardelli, G. Valenza, A. Greco, A. Lanata, and E. P. Scilingo, “Recognizing Emotions Induced by Affective Sounds through Heart Rate Variability,” *IEEE Trans. Affect. Comput.*, vol. 6, no. 4, pp. 385–394, Dec. 2015, doi: 10.1109/TAFFC.2015.2432810.

[5] P. Lang J, Bradley, M.M., and Cuthbert, B.N., “International affective picture system (IAPS): Affective ratings of pictures and instruction manual.,” University of Florida, Gainesville, FL, Technical Report, 2008.

[6] Zheng W.L., Liu W., Lu Y., Lu B.L., and Cichocki A., “EmotionMeter: A Multimodal Framework for Recognizing Human Emotions,” *IEEE Trans. Cybern.*, vol. 49, no. 3, Art. no. 3, 2019, doi: 10.1109/TCYB.2018.2797176.

[7] Gross, J. J., and Levenson, R. W., “Emotion elicitation using films.,” *Cogn Emot*, no. 9, pp. 87–108, 1995, doi: 10.1080/02699939508408966.

[8] A. Chirico, P. Cipresso, D. B. Yaden, F. Biassoni, G. Riva, and A. Gaggioli, “Effectiveness of Immersive Videos in Inducing Awe: An Experimental Study,” *Sci. Rep.*, vol. 7, no. 1, p. 1218, Apr. 2017, doi: 10.1038/s41598-017-01242-0.

[9] J. Marín-Morales *et al.*, “Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors,” *Sci. Rep.*, vol. 8, no. 1, p. 13657, Sep. 2018, doi: 10.1038/s41598-018-32063-4.

[10] C. Wang, H. Zhang, K. S. Phua, T. H. Dat, and C. Guan, “Introduction to NeuroComm: a Platform for Developing Real-Time EEG-based Brain-Computer Interface Applications,” in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2007, pp. 4703–4706, doi: 10.1109/IEMBS.2007.4353389.

[11] A. Gramfort *et al.*, “MEG and EEG data analysis with MNE-Python,” *Front. Neurosci.*, vol. 7, p. 267, 2013, doi: 10.3389/fnins.2013.00267.

[12] A. Seal *et al.*, “An EEG Database and Its Initial Benchmark Emotion Classification Performance,” *Comput. Math. Methods Med.*, vol. 2020, p. 8303465, Aug. 2020, doi: 10.1155/2020/8303465.