AffectivelyVR: Towards VR Personalized Emotion Recognition

Kunal Gupta
kgup421@aucklanduni.ac.nz
The University of Auckland
Auckland, New Zealand

Jovana Lazarevic
jlazarevic.ftn@gmail.com
University of Novi Sad
Novi Sad, Serbia

Yun Suen Pai
yspai1412@gmail.com
The University of Auckland
Auckland, New Zealand

Mark Billinghurst
mark.billinghurst@auckland.ac.nz
The University of Auckland
Auckland, New Zealand

ABSTRACT

We present AffectivelyVR, a personalized real-time emotion recognition system in Virtual Reality (VR) that enables an emotion-adaptive virtual environment. We used off-the-shelf Electroencephalogram (EEG) and Galvanic Skin Response (GSR) physiological sensors to train user-specific machine learning models while exposing users to affective 360° VR videos. Since emotions are largely dependent on interpersonal experiences and expressed in different ways for different people, we personalize the model instead of generalizing it. By doing this, we achieved an emotion recognition rate of 96.5% using the personalized KNN algorithm, and 83.7% using the generalized SVM algorithm.

KEYWORDS

Virtual Reality, Personalized, Emotion Recognition, Physiology, EEG, GSR, Machine Learning

ACM Reference Format:
Kunal Gupta, Jovana Lazarevic, Yun Suen Pai, and Mark Billinghurst. 2020. AffectivelyVR: Towards VR Personalized Emotion Recognition. In *26th ACM Symposium on Virtual Reality Software and Technology (VRST '20), November 1–4, 2020, Virtual Event, Canada.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3385956.3422122

1 INTRODUCTION

In this paper we describe a system for real-personalised emotion recognition in Virtual Reality (VR). Humans are motivated by many different things like fame, fulfilment, happiness, money, or sense of belonging, and by various factors such as instincts, pleasure, obligations, thoughts, and outside forces. However, emotion is a key constant in all of these scenarios, and is a universal part of human nature [6]. Attempts to automatically recognize emotions, such as affective computing [15], has been widely employed using several computational methods on data from correlated input modalities such as face, eye gaze,and physiological signals [4]. Although these methods have been successfully developed using the controlled stimuli in the lab environments, the level of immersion of the set-up has usually been underestimated, thus eliciting emotional experiences that are not similar to the real-world scenarios [1]. Immersive VR can provide the sensation of being in real world and create extremely powerful emotional experiences[2], but there has been relatively little research on how to objectively measure emotional response in VR and project it back into the VR experience [7, 16].

In this work, we present AffectivelyVR, a system that is based on a personalized real-time emotion recognition model that was trained using low-cost physiological sensors. We explore how these sensors can be used to capture signals to measure users' emotional states that can be used to adapt VR elements in real-time. We are aiming for a personalized model as opposed to generalized, meaning the model is trained per participant. Overall, our approach differs from past research in two ways: (1) we present an experimental method to develop the personalized machine learning models to recognize emotional states on real-time basis in VR, and (2) we designed a custom machine learning model architecture with a maximum personalized accuracy of 96.5%.

2 IMPLEMENTATION

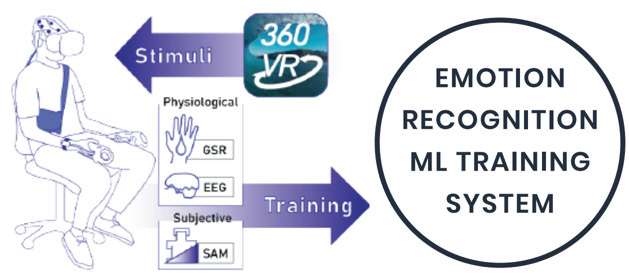

Figure 1: The AffectivelyVR framework, which uses physiological and subjective data to generate machine learning models to predict human emotions in Virtual Reality (VR).

We developed an emotion recognition system using physiological signals, particularly EEG and GSR, while interacting with immersive 360°videos in VR (shown in figure 1). In our system, we used an HTC Vive VR display to enable the participants to experience the VR environment. We collected EEG signals using a 16-channel OpenBCI EEG Cap with daisy chain module at a 125Hz sampling

VRST '20, November 1–4, 2020, Virtual Event, Canada

ACM ISBN 978-1-4503-7619-8/20/11.
https://doi.org/10.1145/3385956.3422122

frequency and focused on the activation of the pre-frontal cortex i.e. FP1 and FP2 [2, 29], occipital, and parietal lobe i.e. O1, O2, P3, Pz, and P4 [28, 40] electrodes in response to the emotion stimulation. A Shimmer sensing device was used to collect GSR signals at a 128Hz sampling rate. We developed a 360° video player application capable of capturing and saving the streamed physiological signals using the Unity Game Engine and LSL4Unity asset while the video was being played.

The emotion recognition model architecture was developed to classify the emotional states from the EEG and GSR data collected during the experimental procedure. At first, the collected raw data was cleaned by eyeballing and removing sudden noise peaks. For EEG pre-processing, we used a band pass filter with low cutoff at 1Hz and high cutoff at 45 Hz followed by a multichannel Wiener filter [17] and Blind Source Separation algorithm i.e. independent component analysis (ICA) to remove EOG, motion and muscle noise artefacts [10]. For GSR pre-processing, the data was first smoothed using a moving average filter and then down-sampled to 4Hz [9]. Finally, the pre-processed data was divided into 5-second epochs with 50% overlap. We used the Biosppy python package [5] for signal processing and extraction of epoch-wise EEG and GSR features consisting of statistical, time-domain, and frequency-domain parameters [8, 13, 18]. In order to label the data, we divided the Self-Assessment Scale (SAM) into two categories, namely Pleasant (valence≥5) and Unpleasant (valence<5).

After the features were extracted, we used the Yeo-Johnson non-linear power transformation [3] that mapped the data from any distribution to as close to a Gaussian distribution as possible in order to stabilize variance and minimize skewness. Then we performed backward feature elimination to remove unnecessary features. Most of the participant data was unbalanced due to inconsistency in the perceived emotional state (SAM ratings reported by participants), so we decided to boost and balance using Principal Component Analysis (PCA) and the Synthetic Minority Oversampling Technique (SMOTE) [14].

From the previous steps, balanced and normalized features were used to train four machine learning classifiers, namely: Gaussian Naive Bayes, k-Nearest Neighbour (KNN), Support Vector Machine (SVM), and Random Forest (RF). We tested the performance of these classifiers with Accuracy, Recall, Precision, and F1 Score as metrics. We used a Grid Search approach to identify the best matching set of hyperparameters for each model with a 10-fold cross-validation. Even though deep learning models are popular for classifying EEG signals [11, 12], they require a lot of training data to reach that level of performance, usually gathered from many participants. In our approach, we are looking towards a personalized model as opposed to generalized, meaning the model is trained per participant.

3 METHOD AND RESULTS

With the goal to develop a personalized emotion recognition system, a within-subject study was conducted where 6 participants (Female:3, Male:3) with a mean age of 34.83 (+/- 14.10 standard deviation) were shown four immersive 360° videos, each 205 seconds long, in a Virtual Reality (VR) environment while recording their physiological responses. As subjective measures, we used the PANAS [38], and SAM [9] self-assessment questionnaires.

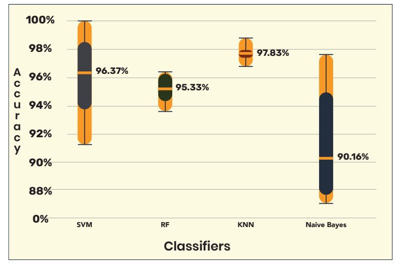

Figure 2: Boxplot of ML Classifiers: SVM, RF, KNN, and Naive Bayes

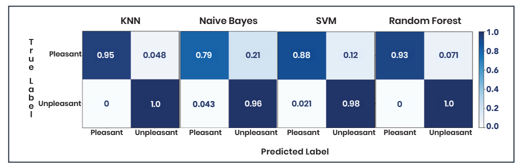

Figure 3: 10-fold Cross-Validated Confusion Matrices for KNN, Naive Bayes, SVM, and RF Classifiers

In AffectivelyVR, all the machine learning classifiers reported more than 85% accuracy for recognizing the participant emotional state, but KNN outperformed SVM, RF and Naive Bayes performed worst of all on an average (figure 2). The 10-fold cross-validated confusion matrices for used classifiers can be seen in figure 3.

4 LIMITATIONS AND FUTURE WORKS

Our aim was to achieve higher accuracy and overcome any possibility of overfitting, so we will collect more physiological data for training and testing personal models by exposing participants to more varied emotional experience trials. We conducted this study with only 6 participants which is very less to provide any statistical inference. So, we plan to increase the number of participants to statistically validate our technique. In the current study, we found it challenging to eyeball all the data and clean sudden motion artefacts. For future works, we aim to automate the motion artefact removal process by using template matching algorithms on the participants' accelerometer data. We also plan to conduct a full user study to evaluate our personalized emotion models to recognize real-time emotions in an emotion adaptive VR application.

5 CONCLUSION

In this work, we presented the AffectivelyVR system, a personalized emotion recognition VR system using players' physiological information to identify their emotional states in real-time. This system aims to assist VR designers and developers to monitor and regulate desired emotional experiences in VR, ultimately improving the overall user experience. We also introduced the machine learning based emotion recognition process for predicting pleasant and unpleasant states using EEG and GSR data, achieving the highest average accuracy of 96.5% and 83.7% accuracy for personalized and generalized models respectively.

REFERENCES

[1] Rosa María Baños, Cristina Botella, Mariano Alcañiz, Víctor Liaño, Belén Guerrero, and Beatriz Rey. 2004. Immersion and emotion: their impact on the sense of presence. *Cyberpsychology & behavior* 7, 6 (2004), 734–741.

[2] Rosa María Baños, Víctor Liaño, Cristina Botella, Mariano Alcañiz, Belén Guerrero, and Beatriz Rey. 2006. Changing induced moods via virtual reality. In *International Conference on Persuasive Technology*. Springer, 7–15.

[3] Anthony J Bishara and James B Hittner. 2012. Testing the significance of a correlation with nonnormal data: comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological methods* 17, 3 (2012), 399.

[4] Rafael A Calvo and Sidney D'Mello. 2010. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on affective computing* 1, 1 (2010), 18–37.

[5] Carlos Carreiras, Ana Priscila Alves, André Lourenço, Filipe Canento, Hugo Silva, Ana Fred, et al. 2018. BioSPPy: Biosignal Processing in Python.

[6] Charles Darwin and Phillip Prodger. 1998. *The expression of the emotions in man and animals*. Oxford University Press, USA.

[7] Arindam Dey, Thammathip Piumsomboon, Youngho Lee, and Mark Billinghurst. 2017. Effects of sharing physiological states of players in a collaborative virtual reality gameplay. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 4045–4056.

[8] Ruo-Nan Duan, Xiao-Wei Wang, and Bao-Liang Lu. 2012. EEG-based emotion recognition in listening music by using support vector machine and linear dynamic system. In *International Conference on Neural Information Processing*. Springer, 468–475.

[9] Kunal Gupta, Ryo Hajika, Yun Suen Pai, Andreas Duenser, Martin Lochner, and Mark Billinghurst. 2019. In AI We Trust: Investigating the Relationship between Biosignals, Trust and Cognitive Load in VR. In *25th ACM Symposium on Virtual Reality Software and Technology*. 1–10.

[10] Kunal Gupta, Ryo Hajika, Yun Suen Pai, Andreas Duenser, Martin Lochner, and Mark Billinghurst. 2020. Measuring Human Trust in a Virtual Assistant using Physiological Sensing in Virtual Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 756–765.

[11] Daekyo Jeong, Sangbong Yoo, and Jang Yun. 2019. Cybersickness analysis with eeg using deep learning algorithms. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 827–835.

[12] Jinwoo Kim, Woojae Kim, Heeseok Oh, Seongmin Lee, and Sanghoon Lee. 2019. A deep cybersickness predictor based on brain signal analysis for virtual reality contents. In *Proceedings of the IEEE International Conference on Computer Vision*. 10580–10589.

[13] Bahareh Nakisa, Mohammad Naim Rastgoo, Dian Tjondronegoro, and Vinod Chandran. 2018. Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors. *Expert Systems with Applications* 93 (2018), 143–155.

[14] Mehdi Naseriparsa and Mohammad Mansour Riahi Kashani. 2014. Combination of PCA with SMOTE resampling to boost the prediction rate in lung cancer dataset. *arXiv preprint arXiv:1403.1949* (2014).

[15] Rosalind W Picard. 2000. *Affective computing*. MIT press.

[16] Thammathip Piumsomboon, Youngho Lee, Gun A Lee, Arindam Dey, and Mark Billinghurst. 2017. Empathic mixed reality: Sharing what you feel and interacting with what you see. In *2017 International Symposium on Ubiquitous Virtual Reality (ISUVR)*. IEEE, 38–41.

[17] Ben Somers, Tom Francart, and Alexander Bertrand. 2018. A generic EEG artifact removal algorithm based on the multi-channel Wiener filter. *Journal of neural engineering* 15, 3 (2018), 036007.

[18] Goran Udovičić, Jurica Đerek, Mladen Russo, and Marjan Sikora. 2017. Wearable emotion recognition system based on GSR and PPG signals. In *Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care*. 53–59.