Multiclass Emotion Classification Using Pupil Size in VR: Tuning Support Vector Machines to Improve Performance

Lim Jia Zheng¹, James Mountstephens², Jason Teo Tze Wi³

¹Graduate Researcher, Evolutionary Computing Laboratory, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia.
²Senior Lecturer, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia.
³Professor, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia.

jtwteo@ums.edu.my

Abstract. Emotion recognition and classification has become a popular topic of research among the area of computer science. In this paper, we present on the emotion classification approach using eye-tracking data solely with machine learning in Virtual Reality (VR). The emotions were classified into four distinct classes according to the Circumplex Model of Affects. The emotional stimuli used for this experiment is 360° videos presented in VR with four sessions stimulation according to the respective quadrant of emotions. Eye-tracking data is recorded using an eye-tracker and pupil diameter was chosen as a single modality feature for this investigation. The classifier used in this experiment was Support Vector Machine (SVM). The best accuracy is obtained from tuning the parameter in SVM and the best accuracy achieved was 57.65%.

1. Introduction

In recent years, the study on emotion recognition has attracted growing interest and it is an increasing trend in the area of computer science. The investigations of emotion detection have become more popular among Human-Computer Interaction (HCI) since the interaction between humans and computers, as well as the role of detect and respond to human emotions. It can be contributed to many domains such as educations, medicine and healthcare, psychology, and computer gaming. There is a study focuses on emotion recognition in HCI [1]. Many contributions of application with emotion detection have been done such as safe driving [2] and emotional security systems [3].

Emotion recognition can be done by using different signals such as non-physiological signals and physiological signals. However, physiological signals are more commonly used by the researchers to classify emotions since it is more reliable. Numerous researches on emotion recognition have been investigated by using different physiological signals such as electrooculograms (EOG) [4], electrocardiogram (ECG) [5], and electroencephalograms (EEG) [6]. There are also research on emotion detection using non-physiological signals such as facial expressions [7] and speech signals [8] but they are more to previous works. There are many emotional investigations conducted by using the combination of multiple physiological signals. However, there is a lack of the usage of eye movement signals solely on the study of emotion detection but it is in an increasing trend.
Machine learning is a data analysis tool that optimizes the creation of analytical models. It can learn from data, recognize patterns and make predictions. There are three types of machine learning algorithms, which are supervised learning, unsupervised learning, and reinforcement learning. In the task of emotion classification, a machine learning algorithm is widely used by the researchers to predict the accuracy and performance of recognition. The common machine learning algorithms include Support Vector Machine (SVM), Random Forest, Decision Tree, and k-Nearest Neighbors (KNNs). However, SVM was used as the classifier for this paper.

Moreover, a stimulation tool is required to evoke the user’s emotions in an emotional experiment. The most common stimulation tool used by researchers including movies and images. However, Virtual Reality (VR) is much less used to present the stimulation. VR is a recent advanced technology that provides a simulated experience in a virtual real-world environment. The user will be fully controlled by the VR headset with the presentation of 360° videos and hence it is fewer distractions in the VR stimuli. The user is immersed in the virtual real-world environment so it has a better experience as well as more real reactions and responses are obtained from the user.

In this paper, we attempt to classify emotions into four distinct classes using eye-tracking data and machine learning in a virtual environment. Eye-tracking data was used as a single modality for this emotional investigation and pupil diameter was chosen as a single feature. The classification tasks were carried out by using Support Vector Machine (SVM) classifier. Emotions were classified and distinguished into four quadrants according to the Circumplex Model of Affects [9]. In the first section of this paper, the introduction of this paper is presented. Section 2 presents the background of emotions, eye-tracking and VR for emotion recognition. Methodology is presented in Section 3 while Section 4 presents the results of the recognition performance with discussion. The last section is the conclusion of this paper.

2. Background

Emotion is described as a state of feeling that correlated with the nervous system and it changes variously with an individual’s thoughts and behavioral actions. Emotions are complex and it is containing various components such as expressive behavioral, verbal behavior, psychophysiological changes, and subjective experience. There is no scientific consensus since different emotional ontologies and theories are used by numerous researchers. Ekman’s model has defined six basic emotions, which are disgust, happiness, sadness, anger, surprise, and fear [10]. Wheel of emotions is then developed in Plutchik’s model with eight basic emotions such as trust, anger, fear, sadness, surprise, joy, disgust, and anticipation [11]. A complex emotion is produced from the relationships between the grouped basic emotions in a positive or negative nature. In emotion classification, Circumplex Model of Affects is a model of emotion that widely used in the study on classifying human emotions. It contains four quadrants with the respective emotions from the result of varying the degree of arousal and valence dimensions.

In emotion recognition, different types of data can be used to classify emotional states such as brainwave signals and eye movement signals. Eye-tracking refers to a method that tracks and measures the eye movements and eye positions. Eye-tracking technology currently is widely used in many domains including computer science and emotional researches. Several eye features that can be utilized to classify emotions such as fixation duration, motion speed of the eyes, pupil positions, pupil dilation and pupil diameter since these eye-tracking data is containing some emotional-relevant features. There are some studies used eye-tracking data as their sensor modality to classify emotions [12-14]. Most of the studies on emotion recognition used the combination of multiple physiological signals such as the combination of EEG signals and eye-tracking data [15], and the combination of EEG and ECG signals [16]. But, it is much less studies that used eye-tracking data solely to recognize emotions.

VR is an artificial reality perceived by sensory stimuli generated by a computer, where one’s behaviors partly decide what is happening in the virtual environment. Within the VR head-mounted display (HMD), the user is fully controlled and the external influences can be dropped to lowest.
Hence, the user is more concentrated to the stimuli and more accurate responses will be obtained. A previous study has shown that Immersive Virtual Environments (IVEs) is potentially be used as a presentation method to trigger emotions [17]. There is a study on affective computing in VR that recognize emotions by using EEG and ECG signals to validate the use of IVEs [18]. There is a recent study on classifying facial expressions using eye-tracking camera in virtual environment [19]. Nowadays, most of the VR HMD is integrated with the eye-tracker hence the eye-tracking data can be easily obtained. Therefore, we adopt this advanced technology with the use of eye-tracking data to classify emotions in a virtual environment.

3. Methods

3.1. Demography target, hardware and experiment setup
This experiment used VR as our stimulation tool. HTC Vive VR headset with a pair of earphones was used to present a 360° video to evoke the emotions of the users. An explanation about the emotional experiment was given to all the participants before the experiment started. The eye-tracking data is recorded simultaneously with the presentation of emotional video clips. The length of the video is about 6 minutes. Ten subjects (10 males) participated in this experiment and the subject’s age range is 21-28. The eye-tracking data is collected using Pupil Labs eye-tracker. The protocol of the experiment is presented in figure 1. There were four sessions of stimulation according to the four quadrants of emotion and there is a 10-seconds rest period after a stimulation session is presented.

![Figure 1. Protocol of the experiment.](image)

3.2. Data collection and classification methods
The eye-tracking data is captured by Pupil Labs software application, namely Pupil Capture. The pupil data is then exported in csv file format from Pupil Player. The exported eye-tracking data included gaze position, pupil diameter, fixation duration, blink, and pupil position. Pupil diameter was chosen as a single feature for this investigation. Python was used as the machine learning language for this experiment. The classifier of this experiment was Support Vector Machine (SVM). The best results will be obtained by tuning the parameters based on the SVM machine learning algorithm.

4. Results and Discussion
In this section, the results obtained are presented in charts and the result comparison of 3 experiments are displayed in the table.
Figure 2. Accuracy comparison of each subject in experiment 1.

Figure 3. Accuracy comparison of each subject in experiment 2.
Three experiments have been conducted in this investigation with SVM by parameter tuning to obtain the best accuracy. In experiment 1, the kernel used is Radial Basis Function (RBF) and the parameter used is gamma, which is used to fit the training datasets. The value is set to [0.1, 1, 10, 100]. In experiment 2, SVM with RBF kernel is used and the gamma value is set to [1, 10, 100, 1000]. While in experiment 3, the RBF kernel with a gamma value of [10, 100, 1000, 2000] is used.

From the results obtained, the highest accuracy obtained from experiment 2 was 57.05% while the highest accuracy in experiment 3 was 56.98%. However, experiment 1 showed the best performance compare to the other two experiments. The highest accuracy achieved was 57.65%. The finding also showed that the results were very close. Most of the results showed that the accuracies were higher when the gamma value is smaller and the accuracies were smallest when there is a higher gamma value.

| Subject ID | Experiment 1 | Experiment 2 | Experiment 3 |
|------------|--------------|--------------|--------------|
| 1          | 45.35%       | 45.73%       | 45.39%       |
| 2          | 36.49%       | 36.16%       | 36.06%       |
| 3          | 42.67%       | 42.56%       | 42.44%       |
| 4          | 37.91%       | 37.61%       | 37.30%       |
| 5          | 53.79%       | 53.50%       | 53.24%       |
| 6          | 57.65%       | 57.05%       | 56.98%       |
| 7          | 39.14%       | 38.81%       | 38.48%       |
| 8          | 49.21%       | 48.96%       | 48.70%       |
| 9          | 43.99%       | 43.52%       | 43.34%       |
| 10         | 39.61%       | 39.04%       | 38.84%       |
5. Conclusion
In this paper, we classified emotions into four distinct classes according to Russell’s Circumplex Model of Affects using machine learning approach. We obtained the eye-tracking data from an eye-tracker and pupil diameter was chosen as the single modality feature for this investigation. The classifier used in this investigation is Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel. By tuning the parameter with the SVM algorithm, the best accuracy was obtained from the three experiments. The findings showed that the highest accuracy was 57.65%. For future work, other eye features will be used to conduct the emotion classification to compare with the performance obtained using pupil diameter.

Acknowledgements
This work was supported by the Ministry of Energy, Science, Technology, Environment and Climate Change (MESTECC), Malaysia [grant number ICF0001-2018].

References
[1] Cowie R, Douglas-Cowie E, Tsapatsoulis N, Votsis G, Kollias S, Fellenz W, Taylor JG. Emotion recognition in human-computer interaction. IEEE Signal processing magazine. 2001 Jan;18(1):32-80.
[2] Alsibai MH, Manap SA. A study on driver fatigue notification systems. ARPN Journal of Engineering and Applied Sciences. 2016;11(18):10987-92.
[3] Cummings EM, Davies P. Emotional security as a regulatory process in normal development and the development of psychopathology. Development and Psychopathology. 1996;8(1):123-39.
[4] Paul S, Banerjee A, Tibarewala DN. Emotional eye movement analysis using electrooculography signal. International Journal of Biomedical Engineering and Technology. 2017;23(1):59-70.
[5] Tivatansakul S, Ohkura M. Emotion recognition using ecg signals with local pattern description methods. International Journal of Affective Engineering. 2015;IJAE-D.
[6] Gupta V, Chopda MD, Pachori RB. Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals. IEEE Sensors Journal. 2018 Nov 27;19(6):2266-74.
[7] Wallbott HG. The robustness of communication of emotion via facial expression: Emotion recognition from photographs with deteriorated pictorial quality. European Journal of Social Psychology. 1991 Jan;21(1):89-98.
[8] Mao Q, Dong M, Huang Z, Zhan Y. Learning salient features for speech emotion recognition using convolutional neural networks. IEEE transactions on multimedia. 2014 Sep 29;16(8):2203-13.
[9] Russell JA. A circumplex model of affect. Journal of personality and social psychology. 1980 Dec;39(6):1161.
[10] Ekman P. Basic emotions. Handbook of cognition and emotion. 1999 Feb 25;98(45-60):16.
[11] Plutchik R. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American scientist. 2001 Jul;89(4):344-50.
[12] Oliva M, Anikin A. Pupil dilation reflects the time course of emotion recognition in human vocalizations. Scientific reports. 2018 Mar 20;8(1):4871.
[13] Lischke A, Berger C, Prehn K, Heinrichs M, Herpertz SC, Domes G. Intranasal oxytocin enhances emotion recognition from dynamic facial expressions and leaves eye-gaze unaffected. Psychoneuroendocrinology. 2012 Apr 1;37(4):475-81.
[14] Rajakumari B, Selvi NS. HCI and eye tracking: Emotion recognition using hidden markov model. International Journal of Computer Science and Network Security (IJCNSNS). 2016 Mar 1;16(3):74.
[15] Zheng WL, Dong BN, Lu BL. Multimodal emotion recognition using EEG and eye tracking...
data. In 2014 36th Annual International Conf. of the IEEE Engineering in Medicine and Biology Society 2014 Aug 26 (pp. 5040-43). IEEE.

[16] Katsigiannis S, Ramzan N. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. IEEE journal of biomedical and health informatics. 2017 Mar 27;22(1):98-107.

[17] Gorini A, Mosso JL, Mosso D, Pineda E, Ruiz NL, Ramiez M, Morales JL, Riva G. Emotional response to virtual reality exposure across different cultures: the role of the attribution process. Cyberpsychology & behavior. 2009 Dec 1;12(6):699-705.

[18] Marín-Morales J, Higuera-Trujillo JL, Greco A, Guixeres J, Llinares C, Scilingo EP, Alcañiz M, Valenza G. Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors. Scientific reports. 2018 Sep 12;8(1):13657.

[19] Hickson S, Dufour N, Sud A, Kwatra V, Essa I. Eyemotion: Classifying facial expressions in VR using eye-tracking cameras. In 2019 IEEE Winter Conference on Applications of Computer Vision (WACV) 2019 Jan 7 (pp. 1626-35). IEEE.