Investigating the Use of Eye Fixation Data for Emotion Classification in VR

Lim Jia Zheng[1], James Mountstephens[2], Jason Teo[3*]

[1]GraduateResearcher, Evolutionary Computing Laboratory, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia
[2]Senior Lecturer, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia
[3]Professor, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia
E-mail: [3*]jtwteo@ums.edu.my

Abstract: Eye-tracking technology has become popular recently and widely used in research on emotion recognition since its usability. In this paper, we presented a preliminary investigation on a novelty approach for detecting emotions using eye-tracking data in virtual reality (VR) to classify 4-quadrant of emotions according to russell'scircumplex model of affects. A presentation of 360^0 videos is used as the experiment stimuli to evoke the emotions of the user in VR. An add-on eye-tracker within the VR headset is used for the recording and collecting device of eye-tracking data. Fixation data is extracted and chosen as the eye feature used in this investigation. The machine learning classifier is support vector machine (SVM) with radial basis function (RBF) kernel. The best classification accuracy achieved is 69.23%. The findings showed that emotion classification using fixation data has promising results in the prediction accuracy from a four-class random classification.

1. Introduction

The study of emotion detection is a fascinating area of study and, with recent developments in high technology, the amount of research has increased significantly. Emotion recognition can be contributed in many domains such as in the field of medicine, education, psychology, and entertainment. Emotions are significant psychological instances that describe many human emotional states. The study of emotional states is useful for both the understanding of human behavior and the integration of human factors into artificial systems. Therefore, the studies focusing on human-computer interaction (HCI)(Lalitha et al., 2015; Jamshidnejad & Jamshidined, 2009) is also in an increasing trend to improve and enhance the interaction between users and the machine based on the understanding of human behavior (Pantic et al., 2007). There are many applications with emotion detection that have been developed and applied in many areas such as marketing (Consoli, 2010) and healthcare (Guo et al., 2013).

Emotion can be detected using non-physiological signals such as face gesture (Gunes & Piccardi, 2007) and speech (Han, Yu, & Tashev, 2014), or physiological signals such as the electrocardiogram (ECG) (Ferdinando, Seppänen, & Alasaarela, 2017), electrooculograms (EOG) (Soundariya & Renuga, 2017), and electromyogram (EMG) (Yang & Yang, 2011), electroencephalograms (EEG) (Qing et al., 2019), and galvanic skin response (GSR) signals (Goshvarpour, Abbasi, & Goshvarpour, 2017). However, compares to non-physiological signals, the scientists are more likely to use physiological signals due to their reliability and usability. EEG signal is the most commonly used since it is a direct connection to the nervous system and human brain activities. Recently, the eye features from eye-tracking data are also widely used by researchers on their emotional investigations. The emotion classification can be done by using several types of eye signals such as pupil size, pupil position, and saccades. A recent report presented the use of eye-tracking for emotion recognition including multiple eye features with current challenges (Lim, Mountstephens, & Teo, 2020).

Moreover, to stimulate an individual's emotional states, an emotion stimulation tool is required. There are several ways to evoke the user's emotions such as watching a film, looking at an image, or listening to music. For example, an emotion can be classified using EEG data by watching a movie (Nie et al., 2011). However, with high technology development, Virtual Reality (VR) is developed. VR is an emerging technology in recent years. Itcan create a 3D virtual world or an immersive environment that is generated by a computer. Using VR as stimuli, the users will be fully controlled and they are immersed in a real-world experience, hence, it is fewer distractions and influences from the outside environment. Therefore, a more realistic response and reaction can be obtained from the user. Furthermore, nowadays VR headset is integrated with eye-tracking technology. It is more convenient and easier for recording and collecting the eye-tracking data when using a VR.

Throughout this paper, we utilized the eye-tracking data to detect emotions and classify them into four distinct classes in VR. In this investigation, fixation data is extracted and used for emotion recognition, and Support Vector Machine (SVM) is the machine learning classifier for classification tasks. Introduction, background, methodology, results, and conclusion are discussed in the topics. This section is briefly including the introduction of emotion recognition. The theory of emotions and eye-tracking with research works are presented in the background section. The methodology section includes the experimental setup, procedure, data collection, and classification method. The results and discussion are presented in the next section. The last section summarizes this paper and discusses future works.

2. Background

Emotions are behavioral states associated with the human brain triggered by neurophysiological changes related to an individual's feelings or behavioral responses. The definition of emotions can also be described as a degree of pleasure or displeasure, it is a positive emotion when happy while it is negative when upset. In emotion classification, it is possible to distinguish between emotional dispositions and emotional episodes. Six basic emotions such as happiness, anger, sadness, surprise, disgust, and fear are examined by Ekman's works (Ekman, 1999). His works are supported by the "wheel of emotions" from Plutchik and two primary emotions are added from the previous six basic emotions, which are trust and anticipation (Plutchik 2000). This emotion wheel depicted the eight fundamental emotions andillustrated various relationships between emotions from the emotion cluster. The Circumplex Model of Affects (Russell 1980), known as the model of emotion classification, is generally used to distinguish emotions. It consists of four quadrants in the dimensional model, which included the valence (positive/negative) and arousal (high/low) dimensions.

Eye-tracking is the method of assessing the eye movements or the focus point ofa user. It is a method of calculating the point of view or the eye's locations and gathering an individual's eye properties and recording them in the form of data. Recently, the research on eye-tracking is increasing significantly since its usability and it can be used in many fields such as neuroinformatic, educations, medical health, or gaming. The eye movement data can be collected by using an eye-tracker or a simple camera and the data collected can be used to detect emotions. The eye properties such as pupil size, pupil position, and fixation duration can be the indicator of emotion recognition. Fixation is the preservation of an individual's visual gaze in a single position. Several features can be extracted from fixation data such as the number of fixation counts, fixation duration, and first fixation. There are studies on emotion recognition using eye-tracking data such as pupil size(Lanatà et al., 2011) and gaze position (Aracena et al., 2016).

VR is one of the recent advanced technology that can create a simulated environment that is generated by a computer program. The user has immersed in a virtual world within the VR headset hence a distraction from the outside environment can be dropped to the lowest. Therefore, an authentic reaction and response can be obtained from the user in VR stimuli. It is a good stimulation method to elicit the emotions of the user. VR nowadays is integrated with eye-tracking technology so it is convenient and easy for obtaining eye-tracking data. Thus, the objective of this paper is to investigate the use of fixation data on emotion recognition based on the four distinct quadrants from Russell's emotion modelin VR stimuli using an SVM machine learning classifier.

3. Methods

3.1 Experiment setup

In this emotional experiment, VR is used as our stimulation tool. A 360^0 video is presented using an HTC Vive VR headset with a pair of earphones to stimulate the emotions of the user. The contents of the emotional video clips are presented in four sessions according to the four distinct classesbased on thevalence and arousal dimension as shown in Figure 1.
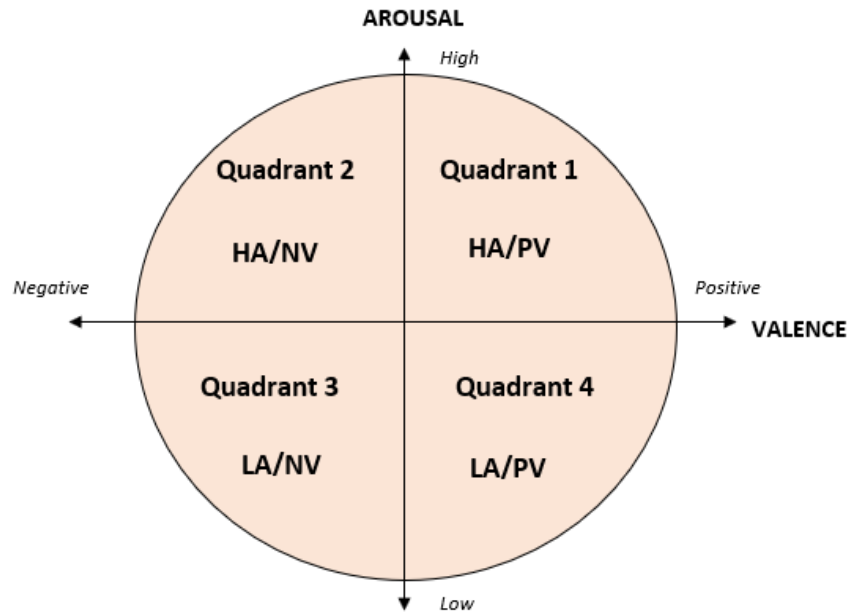
Figure1. Illustration of valence/arousal dimensional model

3.2 Procedure

There was a total of 10 subjects (10 males) volunteered in this investigation and the age range is 22-29. Every participant was given an explanation and procedure of the investigation before the experiment began. For the recording and collecting of eye-tracking data, the hardware used is an add-on Pupil Labs eye-tracker, which is attached inside the VR headset. Figure 2 illustrated the flow of the video presentation when conducting the experiment. The total duration of the video presentation is approximately 6 minutes. It was given a 5-seconds preparation for VRstartup before the video clips started. There was a 10-second resting time between every video session until the video ended and it was about 80 seconds for every stimulation session.
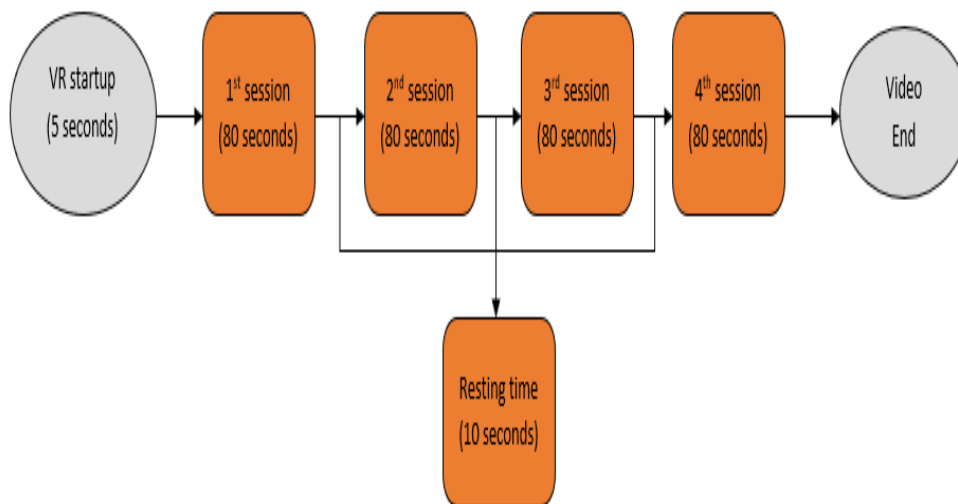

Figure2. Flow of the video presentation

Figure 3 shows the VR stimuli used by the participants in the experiment. The participant sitsin front of the computer screen monitor. They are allowed to turn around their head and body in 360 degrees at their sitting position when watching the video. The participant is not allowed to stand up or move with a big step due to the limited wired attached to the VR headset.

Figure 3.VR headset used by the subjects in the experiment

3.3 Data collection and classifications

In data collection, the recording and capturing of eye-tracking data were done by using Pupil Capture, software with the Pupil Core Headset. The collected data is then undergoing the process of data visualization and exportation by using Pupil Player. The data is saved in CSV file format. The classification is done by using the fixation data. The Support Vector Machine (SVM) machine learning algorithm is used for the emotion classification tasks with the kernel of the Radial Basis Function (RBF). The running script is prepared by using Python programming language.

4. Results and Discussion

In this section, the following are the results obtained from the emotion classification. They are presented in the chart and table.
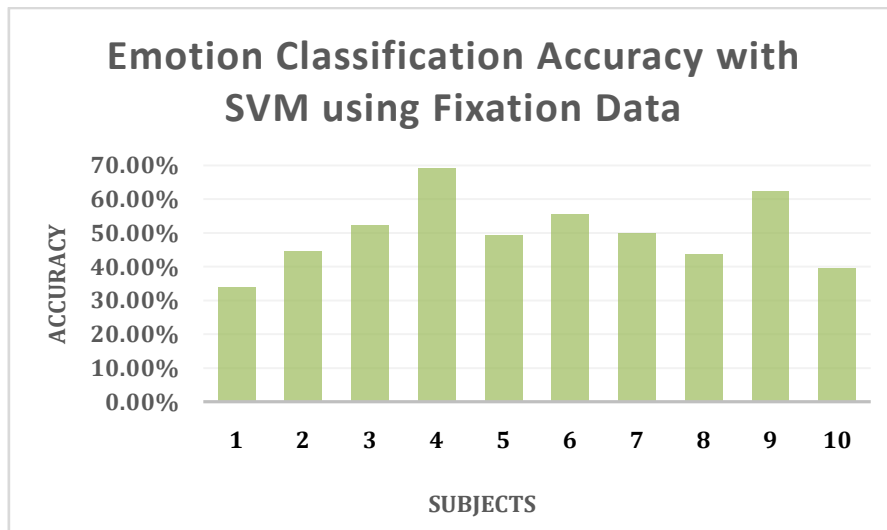

Figure4. Classification accuracy for the 10 subjects in the experiment

Table 1. Tabulated result of emotion classification

Subject	Classification Accuracy
1	33.84%
2	44.71%
3	52.30%
4	69.23%
5	49.47%
6	55.55%

7	48.97%
8	43.60%
9	62.50%
10	47.76%

All classification results were obtained from the SVM machine learning algorithm with RBF kernel for each of the subjects in the experiment. Firstly, the user's eye data is recorded simultaneously when they are watching the video stimuli. The eye data is extracted in the form of data with the timestamp. The collected eye data is then matched up with video presentation time according to each of the session's quadrant. The unused data columns are then be removed. The data is saved and imported to the SVM machine learning algorithm to run the classification process.

The classification accuracy of each subject is presented in the chart and table for comparison. Fixation data is used as a single modality for emotion classification. We extracted the position of the fixation's centroid to classify the emotions. According to the results, the highest accuracy obtained was 69.23% while the 33.84% accuracy was the lowest performance obtained. The findings showed that emotion classification using fixation data has promising results in the prediction accuracy that close to 70% from a four-class random classification. Compares to the previous study (Zheng, Mountstephens, and Teo 2020) that using pupil diameter to classify emotions, emotion recognition using fixation data has a better performance.

5. Conclusion and Future Work

The main purpose of this paper is to investigate the use of fixation data for emotion classification based on the four distinct quadrants from Russell's emotion model in VR stimuli. We extracted the fixation data and uses it as the single emotional-relevant eye feature in this investigation. The machine learning classifier used in the experiment is SVM with RBF kernel and the highest accuracy obtained was 69.23% from this machine learning algorithm. The findings show that the results are promising in emotion recognition since the accuracy achieved at almost 70% from the random classifier. For future work, more features from the fixation data will be extracted and used for the emotion classification in a future experiment such as the fixation duration and total distance traveled from the fixation's position to compare the output here.

6. Acknowledgements

References

1. Aracena, Claudio, Sebastian Basterrech, Vaclav Snasel, and Juan Velasquez. 2016. "Neural Networks for Emotion Recognition Based on Eye Tracking Data." Proceedings - 2015 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015, 2632–37. https://doi.org/10.1109/SMC.2015.460.

2. Consoli, Domenico. 2010. "A New Concept of Marketing: The Emotional Marketing." Brand Research in Accounting, Negociation and Distribution 1 (1): 52–59.

3. Ekman, Paul. 1999. "Basic Emotions." Encyclopedia of Personality and Individual Differences. https://doi.org/10.1007/978-3-319-28099-8_495-1.

4. Ferdinando, H., T. Seppänen, and Esko Alasaarela. 2017. "Enhancing Emotion Recognition from ECG Signals Using Supervised Dimensionality Reduction." ICPRAM 2017 - Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods 2017-Janua (Icpram): 112–28. https://doi.org/10.5220/0006147801120118.

5. Goshvarpour, Atefeh, Ataollah Abbasi, and Ateke Goshvarpour. 2017. "An Accurate Emotion Recognition System Using ECG and GSR Signals and Matching Pursuit Method." Biomedical Journal 40 (6): 355–68. https://doi.org/10.1016/j.bj.2017.11.001.

6. Gunes, Hatice, and Massimo Piccardi. 2007. "Bi-Modal Emotion Recognition from Expressive Face and Body Gestures." Journal of Network and Computer Applications 30 (4): 1334–45. https://doi.org/10.1016/j.jnca.2006.09.007.

7. Guo, Rui, Shuangjiang Li, Li He, Wei Gao, Hairong Qi, and Gina Owens. 2013. "Pervasive and Unobtrusive Emotion Sensing for Human Mental Health." Proceedings of the 2013 7th International

Conference on Pervasive Computing Technologies for Healthcare and Workshops, PervasiveHealth 2013, 436–39. https://doi.org/10.4108/icst.pervasivehealth.2013.252133.

8. Han, Kun, Dong Yu, and Ivan Tashev. 2014. "Speech Emotion Recognition Using Deep Neural Network and Extreme Learning Machine." Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, no. September: 223–27.

9. Jamshidnejad, Amir, and A. Jamshidined. 2009. "Facial Emotion Recognition for Human Computer Interaction Using a Fuzzy Model in the E-Business." 2009 Innovative Technologies in Intelligent Systems and Industrial Applications, CITISIA 2009, no. July: 202–4. https://doi.org/10.1109/CITISIA.2009.5224214.

10. Lalitha, S., Sahruday Patnaik, T. H. Arvind, Vivek Madhusudhan, and Shikha Tripathi. 2015. "Emotion Recognition through Speech Signal for Human-Computer Interaction." Proceedings - 2014 5th International Symposium on Electronic System Design, ISED 2014, 217–18. https://doi.org/10.1109/ISED.2014.54.

11. Lanatà, Antonio, Antonino Armato, Gaetano Valenza, and Enzo Pasquale Scilingo. 2011. "Eye Tracking and Pupil Size Variation as Response to Affective Stimuli: A Preliminary Study." 2011 5th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, PervasiveHealth 2011, 78–84. https://doi.org/10.4108/icst.pervasivehealth.2011.246056.

12. Lim, Jia Zheng, James Mountstephens, and Jason Teo. 2020. "Emotion Recognition Using Eye-Tracking: Taxonomy, Review and Current Challenges." Sensors (Switzerland) 20 (8): 1–21. https://doi.org/10.3390/s20082384.

13. Nie, Dan, Xiao Wei Wang, Li Chen Shi, and Bao Liang Lu. 2011. "EEG-Based Emotion Recognition during Watching Movies." 2011 5th International IEEE/EMBS Conference on Neural Engineering, NER 2011, 667–70. https://doi.org/10.1109/NER.2011.5910636.

14. Pantic, Maja, Alex Pentland, Anton Nijholt, and Thomas S. Huang. 2007. "Human Computing and Machine Understanding of Human Behavior: A Survey." Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 4451 LNAI: 47–71. https://doi.org/10.1007/978-3-540-72348-6_3.

15. Plutchik, Robert. 2000. Emotions in the Practice of Psychotherapy: Clinical Implications of Affect Theories.Emotions in the Practice of Psychotherapy: Clinical Implications of Affect Theories. Washington, DC, US: American Psychological Association. https://doi.org/10.1037/10366-000.

16. Qing, Chunmei, Rui Qiao, Xiangmin Xu, and Yongqiang Cheng. 2019. "Interpretable Emotion Recognition Using EEG Signals." IEEE Access 7: 94160–70. https://doi.org/10.1109/ACCESS.2019.2928691.

17. Russell, James A. 1980. "A Circumplex Model of Affect." Journal of Personality and Social Psychology 39 (6): 1161–78. https://doi.org/10.1037/h0077714.

18. Soundariya, R. S., and R. Renuga. 2017. "Eye Movement Based Emotion Recognition Using Electrooculography." 2017 Innovations in Power and Advanced Computing Technologies, i-PACT 2017 2017-Janua: 1–5. https://doi.org/10.1109/IPACT.2017.8245212.

19. Yang, Guangying, and Shanxiao Yang. 2011. "Study of Emotion Recognition Based on Surface Electromyography and Improved Least Squares Support Vector Machine." Journal of Computers 6 (8): 1707–14. https://doi.org/10.4304/jcp.6.8.1707-1714.

20. Zheng, Lim Jia, James Mountstephens, and Jason Teo. 2020. "Comparing Eye-Tracking versus EEG Features for Four-Class Emotion Classification in VR Predictive Analytics." International Journal of Advanced Science and Technology 29 (6 Special Issue): 1492–97.