Exploring Pupil Position as An Eye-Tracking Feature for Four-Class Emotion Classification In VR

Jia Zheng Lim¹, James Mountstephens² and Jason Teo³

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.

Email: limjiazheng94@gmail.com

Abstract. This paper presented a preliminary investigation of a novel approach on emotion recognition using pupil position in Virtual Reality (VR). We explore pupil position as an eye-tracking feature for four-class emotion classification according to the four-quadrant model of emotions via a presentation of 360° videos in VR. A total of ten subjects participated in this emotional experiment. A 360° video with four sessions of stimulation of emotions will be presented in VR to evoke the user’s emotions. The eye data were recorded and collected using Pupil Labs eye-tracker and the emotion classification was done by using pupil position solely. The classifier used in this investigation is the Support Vector Machine (SVM) machine learning algorithm. The results showed that the best accuracy achieved from this four-class random classification was 59.19%.

1. Introduction

Recently, the research on emotion recognition using eye-tracking is in an increasing trend in computer science especially affective science due to the usability and reliability. In affective science, emotion classification refers to a process, which differentiates one emotion from another. In understanding the feelings or emotions of others, human differs greatly in their precision, the use of emotion recognition system helps to assist individuals with emotion detection. Some emotion detection applications have been proposed in real-life such as the human healthcare system [1] and emotional marketing [2].

There are two types of signals that can be used to classify emotions, which are physiological signals and non-physiological signals. Physiological signals refer to the reading and measurement results from a sensor or device provided by human beings’ physiological processes. Several information can be obtained for emotion classification such as pupil signals, brainwave signals, and heartbeat rate. Many emotional studies have been published by using various type of physiological signals such as galvanic skin response (GSR) signals [3], electrocardiogram (ECG) [4], electromyogram (EMG) [5], and electroencephalograms (EEG) [6]. Besides that, eye movement signals and gaze information can be applied for emotion recognition such as pupil width, gaze positions, and eye fixations. While non-physiological signals include human physical signals such as body gestures [7], speech signals [8], and facial expressions [9].
Furthermore, several techniques or tools can be used by investigators to evoke the emotions of their users. Some popular stimulation tools included watching a video clip or movie, or listening to music. Virtual reality (VR) refers to an advanced technology that creates a simulated environment using a computer program in which a user can experience and interact in a 3-dimensional (3D) world. The user is completely controlled from a VR view, such as the hearing and sight senses within the VR headset. Moreover, most of the VR is currently made with the integration of eye-tracking technology, therefore the eye data can be easily captured and collected from the user.

In this study, we explore pupil position as an eye-tracking feature for four-class emotion classification according to the four-quadrant model of emotions via a presentation of 360° videos in VR. A brief introduction is presented in the first section of this study. The background including related works regarding our topic is explained in Section 2. The next section is the methodology parts, while Section 4 presented the results and discussion. The last part is the conclusion of this report.

2. Background

Emotion is described as a mental reaction that is subjectively perceived as a strong sensation that is usually directed towards a specific object and typically accompanied by cognitive and physiological changes [10]. The theory of discrete emotion, it is started with six fundamental emotions, which are fear, anger, happiness, disgust, surprise, and sadness, which is proposed by Ekman’s model [11]. His work was then supported by Plutchik’s model of emotion, which introduced eight primary emotions and added on two emotions: trust and anticipation, into the previous six emotions from Ekman [12]. The wheel of emotions indicated these eight emotions categorized on a positive/negative perspective. Emotion classification refers to the distinction between the complex emotions associated with the degree level of pleasure or displeasure. The Circumplex model from Russell distributed the emotions into an arousal/valence dimension [13]. The vertical axis (arousal) has a high/low variation, while the horizontal axis (valence) has a positive/negative variation. Four distinct quadrants are generated in a two-dimensional circular space and the respective emotions will be produced in each of the quadrants based on the level of the two dimensions.

Eye-tracking technology allows a computer or machine to know where an individual is looking, whereas an eye-tracker is the sensor device to measure the eye positions and eye properties. Several eye features can be extracted from the eye movement signals for emotion classification such as pupil size, fixation duration, and pupil position. The popular eye characteristics such as pupil diameter [14] and fixation [15] are commonly used by investigators to conduct their emotional research. Also, some researchers used the potential difference of electrodes, called electrooculography (EOG) to conduct their emotional experiment [16]. However, the study on emotion classification relies on pupil position solely is limited. There is a neural network approach for emotion recognition using pupil size and gaze position [17]. Hence, this paper proposed an approach for emotion recognition that relies on pupil position alone in VR.

VR is a computer-generated world with a realistic environment or objects that allow the user to experience the immersive surrounding environment. In an emotional experiment, VR allows the user to immerse themselves in the presentation of 360° videos inside a VR helmet, called head-mounted display (HMD). A report has shown that Immersive Virtual Environments (IVEs) is potentially be used as a good stimulation tool for stimulating emotions [18]. By using VR stimuli, it is fewer distraction and influences from the outside surrounding when the user is wearing the VR headset. It also can effectively narrow down the inaccuracy for obtaining and collecting the user’s emotions.

3. Methodology

3.1 Stimuli, subjects, and experiment procedure

In our experiment, VR with a presentation of 360° videos is used to stimulate the user’s emotions. The hardware included the HTC Vive VR headset, Pupil Labs eye-tracker, and a set of a high-end graphics computer program. The video contains four quadrants of emotions based on the Circumplex model.
Ten subjects (5 males and 5 females) with ages ranging from 21-28 took part in this experiment. All participants are healthy and had normal hearing and vision. Each of the participants will be given an explanation of the experimental protocol before the experiment begins. Before the experiment starts, the calibration process is planned and ready. The participant is allowed to sit in front of the monitor and wear the VR headset. Due to the limited length of a wired connection to the VR headset, participants are only allowed to sit down but they can move around their head and body on the chair in 360° when the video is presenting. There were four separate sessions in the video and each of the sessions represented the respective quadrant of emotion. Each session is about 80 seconds and a 10-seconds resting period will be given before the next session started. The total duration of the video is about 6 minutes. Figure 1 presented the protocol of the experiment.

![Figure 1. Illustration of the experimental protocol.](image)

### 3.2 Data collection and classification

In the experiment, Pupil labs eye-tracker is used for the eye data capturing and recording. The pupil information is gathered and collected using Pupil Capture. The data visualization and exportation were done by using Pupil Player. The raw data exporter from the Pupil Labs framework extracts and exports the collected pupil information to a csv file format and is active by default. Pupil position was selected and explored as an eye feature for this emotional experiment. A classifier of Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel was used to perform the classification tasks. A python programming script of the SVM machine learning algorithm was used to run the classification.

### 4. Results and discussion

The outcomes of emotion classification using the SVM classifier are presented through the chart and table.

![Figure 2. Comparison of emotion classification results for 10 subjects in the experiment.](image)
Table 1. Tabulated results of emotion classification.

| Subject | Emotion Classification Accuracy |
|---------|---------------------------------|
| 1       | 49.42%                          |
| 2       | 57.85%                          |
| 3       | 59.19%                          |
| 4       | 37.62%                          |
| 5       | 48.92%                          |
| 6       | 48.20%                          |
| 7       | 52.03%                          |
| 8       | 46.45%                          |
| 9       | 49.29%                          |
| 10      | 46.19%                          |

The classification tasks are done by using an SVM classifier with RBF kernel to obtain the best emotion recognition accuracy. All the classification was done by using pupil position solely without the combination of other eye features. Figure 2 and table 1 displayed the comparison of emotion classification results for 10 subjects in the experiment. From the chart and table, the best performance achieved was the accuracy of 59.19%, while the lowest accuracy was 37.62%. The highest accuracy showed a promising result that close to 60% from a four-class random classification. The findings showed that pupil position is potentially be used as a single feature for emotion recognition.

5. Conclusion

This paper presented a preliminary investigation of an approach on emotion classification using pupil position. The main purpose of this study is to explore pupil position as an eye feature for four-class emotion recognition based on the Circumplex model of emotions in VR. A machine learning of SVM with RBF kernel is used for the classification tasks. The highest accuracy achieved was 59.19%, which is close to 60% from a random classification. This finding shows that pupil position from eye-tracking data is potentially be used as an emotional feature for emotion recognition. For future work, the pupil position-based emotion classification experiment will be done by using machine learning with parameter tuning to obtain a better performance and compare the results obtained in this paper.

References

[1] Guo R, Li S, He L, Gao W, Qi H and Owens G 2013 Int. Conf. Pervasive Comput. Technol. Healthc. Work. PervasiveHealth pp. 436–439.
[2] Consoli D 2010 Brand Res. Accounting, Negoc. Distrib. 1 52–59.
[3] Wu G, Liu G and Hao M 2010 Int. Symp. Intell. Inf. Process. Trust. Comput. pp. 360–363.
[4] TIVATANSAKUL S and OHKURA M Int. J. Affect. Eng 15 51–61.
[5] Wei C Z 2017 Adv. Mater. Res. 709 827–831.
[6] Alhagry S, Aly A and R A 2017 Int. J. Adv. Comput. Sci. Appl. 8 10.
[7] Glowinski D, Camurri A, Volpe G, Dael N and Scherer K 2008 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work. pp 20.
[8] Liu J, Han W, Ruan H, Chen X, Jiang D and Li H 2018 1st Asian Conf. Affect. Comput. Intell. Interact. ACII Asia 16 2203–2213.
[9] Zhang Y D 2016 IEEE Access 4 8375–8385.
[10] Cabanac M and Cabanac M 2016 What is emotion? What is emotion? 6357 69–83.
[11] Ekman P 1999 Encyclopedia of Personality and Individual Differences 1–6.
[12] R. Plutchik 2001 Philos. Stud. 52 393–409.
[13] Russell J A 1980 J. Pers. Soc. Psychol. 39 1161–1178.
[14] Zheng L J, Mountstevens J and Teo J 2020 J. Big Data 7 1.
[15] Lischke A, Berger C, Prehn K, Heinrichs M, Herpertz S C and Domes G 2012 Psychoneuroendocrinology 37 475–481.
[16] Soundariya R S and Renuga R 2017 Innov. Power Adv. Comput. Technol. i-PACT pp. 1–5.
[17] Aracena C, Basterrech S, Snasel V and Velasquez J 2016 Int. Conf. Syst. Man, Cybern. SMC pp. 2632–2637.
[18] Gorini A 2009 CyberPsychology Behav. 12 699–705.

Acknowledgements
This work was supported by Ministry of Science, Technology and Innovation (MOSTI), Malaysia [grant number ICF0001-2018].