Pilot Study of Emotion Recognition through Facial Expression

Radin Puteri Hazimah Radin Monawir[1], Noraidah Blar[1], Fairul Azni Jafar[2], Zulhasnizam Hasan[3]
[1]Center of Graduate Studies, Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Melaka, Malaysia.
[2]Department of Robotics and Automation, Faculty of Manufacturing Engineering,
Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Melaka, Malaysia.
[3]Department of Electronics and Computer Engineering, Faculty of Engineering Technology,
Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Melaka, Malaysia.
radinhazimah90@gmail.com

***Abstract*— This paper presents our finding from a pilot study on human reaction through facial expression as well as brainwave changes when being induced by audio-visual stimuli while using the Emotiv Epoc equipment. We hypothesize that Emotiv Epoc capable to detect the emotion of the participants and the graphs would match with facial expression display. In this study, four healthy men were chosen and being induced with eight videos, six videos are predefined whereas the other two videos are personalized. We aim for identifying the optimum set up for the real experiment, to validate the capability of the Emotiv Epoc and to obtain spontaneous facial expression database. Thus, from the pilot study, the principal result shows that emotion is better if being induced by using personalized videos. Not only that, it also shows the brainwave produced by Emotiv Epoc is aligned with the facial expression especially for positive emotion cases. Hence, it is possible to obtain spontaneous database in the present of Emotiv Epoc.**

***Index Terms*— Emotion recognition; Facial expression, Brainwave changes.**

I. Introduction

Emotion is usually described as feelings that have been translated physically or biologically by the human body. Scientifically, emotion is a result of brain activities which lead to either voluntary or involuntary reactions of the human body, such as smiling, increase in heartbeat, and yelling. In a general way, these reactions can be simply referred as body gesture, biological signal, facial expression, and speech. These reactions have significant characteristics that directly related to the emotions that a person want to express.

Basically, most researches commonly study facial expression, biological signal, and speech for classifying emotions. Emotion recognition by using facial expression is quite a popular study since it has visible features, is frequently used to convey feelings, and has a great potential to be commercialized. This method used distinct features on the faces such as lips, eyes and eyes brows which also known as the Action Unit (AU) to recognize emotion. Whereas, biological signals are often used for recognizing emotion in medical research. Physiological features, for instance, heartbeat, heart rate, Galvanic skin response, blood volume and pressure, respiration pattern and brainwave signals are used for indicating emotional and psychological health [1],[2]. Meanwhile, emotion recognition via speech normally involves acoustic-prosodic and lexical characteristics [3-6]. There are also hybrid approaches which combine two or more signals in order to classify the emotion, for examples speech and visual fusion and multimodal biological signal. Even hybrid system is usually more accurate and robust than single analysis, it also more time consuming and complex than single analysis. On the other hand, single analysis by using speech has accurateness up to 60%, while by using facial expression can convey emotion accurateness up until 70-98% [2]. However, both physical appearance and voice can be deceived, so, obtaining natural reaction as database for emotion recognition is a challenge in many studies.

In terms of applications, intelligent emotion recognition system is very beneficial to many fields, especially the one which involve interaction of human with robots or computers. For instance, in the education field, Litman's research team use speech features to recognize student emotion during tutoring session [7],[8]. In other research, an educational game was developed by a group of Singaporean and Chinese researchers in order to understand students' interest by looking into their achievement during the session [9]. The information from the Fuzzy game will be used for improving learning session to suit the student interest later. Not only that, many studies have been done by using biological signals for tracing current emotion or psychological state of patients. Therefore, this physiological data have been used by many medical researchers for diagnosing many psychological illnesses such as schizophrenia [10], depression [11] and anxiety [12].

In our project, we would like to study an emotion recognition system based on spontaneous facial expression since we seek for a balance between a fast computing and high accuracy system. However, since facial expression and self-reported answers can be deceived, a brainwave detector is used to check the authenticity of the expression and survey answers. Therefore, in this initial study, an experiment will be carried out with Emotiv Epoc, as evaluator of the emotion display. Emotiv Epoc is a low cost neurohead set that used 14 sensors to detect brainwave. This tool has advantages over other methods of brainwave study such as magnetoencephalography (MEG), functional magnetic resonance imaging (fMIR) and positron emission tomography (PET), since it enables real time brain activity detection, low cost implementation and able to work under various surrounding [13]. In this initial study, we hypothesized that the brainwave which is detected by the Emotive Epoc should be matched with facial expression as well self-answered report.

Therefore, this paper's purpose is threefold which are to figure out suitable method and design of experiment to obtain spontaneous facial expression, to prove the reliability of the Emotiv Epoc to detect the brain signals according to the emotion as well as to collect genuine facial expression images as data base training. Next, we treated emotion as two general groups which are positive and negative emotions. In Section 2, data collection and experiment set up are discussed. The results of emotion and expression display will be discussed in Section 3, followed by conclusion in Section 4.

II. METHODOLOGY

A. Experimental Set Up

The experiment set up for gathering facial expression and brainwave is as in Figure 1. The equipments are brainwave sensor, earphones, Android smartphone camera including its holder, and two laptops; one laptop for inducing emotion while another laptop for monitoring and recording the reaction of the participants.

The Android smartphone is connected to the monitoring laptop via Bluetooth. An Android application called SmartCam which enables smartphone to be a temporary webcam to the laptop is installed in both laptop and the smartphone. The advantages of using the smartphone camera over other acquisition tools (for instance digital camera and webcam) are it enables real time data transferred to laptop, portable and easy to be connected to any laptop as well as can be adjusted to get the suitable angle. It is also affordable yet almost-all-people have smartphone. Moreover, the resolution of the image can be set according to smartphone specification. Therefore, the better the resolution that smartphone can support, the better the image quality would be.

Meanwhile, Emotiv Epoc headset is a headset that is able to detect EEG signals of the brain. The Emotiv Epoc enables the detection of brainwave easier since it has simple build up and user-friendly interface. Therefore, mere participants can use it themselves without specific training. The electrodes of the Emotiv Epoc need to be damped with saline solution in order to get the brainwave. Beside SmartCam and Emotiv Epoc software, we also use open source software for laptop display recording which called CamStudio. Figure 2 shows the overview of the software and hardware involved.

B. Experimental Procedure

As mention earlier, we concluded the emotions into two universal groups which are positive and negative emotions. For exciting the emotions, we constrained the audio-visual stimulants into eight YouTube videos; six of them are predefined while another two are the experiment subject's own-preferred videos. The ratio between positive and negative emotion-induced videos is balanced. Positive emotions are aroused through funny and cheerful videos. Meanwhile, negative emotions are triggered by using sad and angry sentiment. Noted, the duration of each video is limited to less than six minutes to avoid boredom. The sources of the videos are stated in the references. Next, videos are presented to subjects using Window media player while Window 7 is used to browse the questionnaires.

Since we want to obtain spontaneous reaction of the candidates, they need to wear Emotiv Epoc while watching the videos and answering the questionnaire (as in Figure 3). The purpose of the survey is to investigate whether the candidate can feel the emotion or not by stating their emotion before, during, and after watching each video. They also need to state their opinion about the real emotion the video try to convey, despite of their real emotion while watching it. On the other hand, Emotiv Epoc acts as an aid for double checking the truth of the questionnaire and expression. It can be done by matching the brainwave with those results.

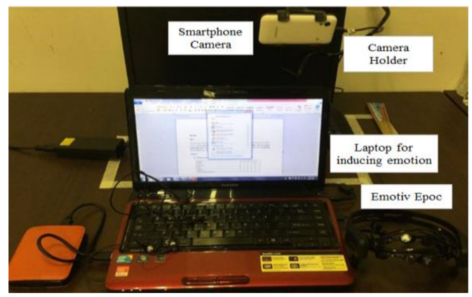

Figure 1: Hardware set up

Figure 2: CamStudio interface (upper left), SmartCam interface (upper middle), Emotiv Epoc headset (upper right), and desktop of monitoring laptop display (below).

The participants in this experiment are four healthy men who are fluent in English. Health is the main consideration since Emotiv Epoc cannot be worn by someone with heart problems. Next, we decided to do experiment only on man in our pilot study since it is easier for Emotiv Epoc electrodes to detect scalp of short-hair man rather than woman. This is because woman normally has thick and long hair which would probably make the detection become difficult. Note that the experiment need to be carried out in a closed environment with good illumination and temperature is controlled at room temperature, which is 27 to 29 degree Celsius. The sequence of the experiment is as in Figure 4. All eight videos were displayed to the participants simultaneously with positive and negative emotion induced videos mixed randomly.

Figure 3: Preliminary experiment set up. The android smartphone camera is fixed above the laptop in appropriate length so that frontal facial expression can be recorded. Meanwhile, the Emotiv Epoc is placed at the head of the participants.

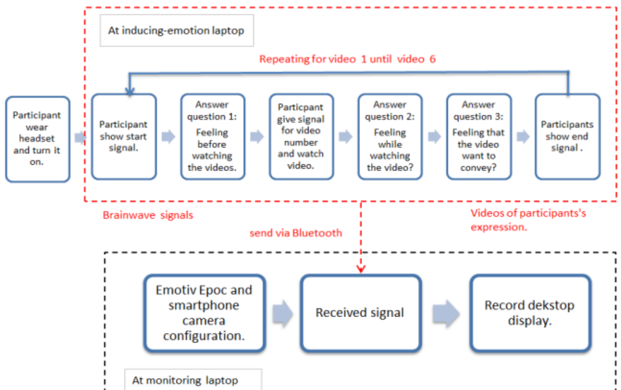

Figure 4: Flow chart of experiment sequences. The questionnaires only need to be answered while watching videos 1 until 6.

C. Data Collection

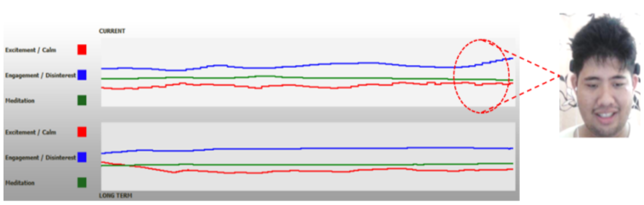

Figure 5: Graph produced by Emotiv Epoc. Upper graph shows current brainwave of the participant, meanwhile below graph shows participant's brainwave for a long period of time.

Data collection is done by extracting and matching the brainwave with facial expression display manually. During the experiment (refer Figure 4), CamStudio software record the desktop display of the monitoring laptop (as in Figure 2), which later been used for data extraction (refer Figure 5).

As we can see in Figure 5, there are two graphs, the upper graph represents current emotion feels by the participant, whereas the lower graph indicates the emotion of the participant for a long period over the whole experiment time. Next, the three coloured lines represent the mood of the participants. Red line is labeled as excitement and calm emotion, the blue line represents engagement and disinterest, and the green line indicates meditation. In other words, red line can be described as emotion detector, the blue line portrays the depth of participant attention and finally the green line is defined as the natural state of the participant. Therefore, for this experiment only red line is significant to ensure that the expression is complied with the brainwave.

III. Results and Discussion

In this experiment, we induced the participants' emotions by using eight videos where six of them are already selected earlier together with the questionnaires. However, the other two videos which contain both negative and positive emotions are selected by the participants themselves and no questionnaire is required to be answered. We deduce that questionnaire is only needed for the videos that are preselected since the videos might not induce emotion as we expected. However, for participants' video selection, we assume that the chosen videos would surely induce the emotion as required since it suits their preference. Actually, the idea of having extra two own preferred videos come after looking the initial results of the first participant.

During the experiment of the first participant (participant A), he did not show much interest and reaction both in brainwave graph and facial expression (referred Figure 6 and Figure 7). At the end of the session, we found that he already watched the videos. This is the main reason he did not feel excited like he was before. Therefore, considering that the same situation might happen to the other participants, we decided to take precaution by having each participant to select their own videos.From Figure 6, the red line should be excited above the green line when positive emotion is induced. However, from all the current emotion graph, only the last graph shows significant inclination. As for facial expression, even though first graph shows smiling face, the brainwave does not incline above the meditation line. Besides, the smile in the first photo is less cheerful as compared to the last photo. This proves that video which is selected based on a person preference can best excite the positive emotion.

In contrast to positive emotion, negative emotion is hard to be detected by using current emotion graphs. As we can see in Figure 7, the red line in current graphs only moves around the meditation line. Therefore, instead of using the current emotion graph, negative emotion is portrayed through the long term graph. From Figure 7, most red lines show declination in the long term graph. As for facial expression, the expressions displayed while watching all negative stimuli do not show so much different from one another.

As the result of having participants' selection video, most of the participants show more interest in the personalized-videos, results in significant changes in brainwave as well as the expressed emotion, just like participant A. Next, Emotiv Epoc graph changes solely depending on the feeling of the wearer. Let say if the wearer just pretends to be happy and smile broadly, the graph would not show an inclination in red line as it should be (as in Figure 8). Interestingly, if we look into detail, the features of mimicking happy face also different from the face features of spontaneous expression. Features of the mimicking face seem to have greater intensity as compared to spontaneous face.

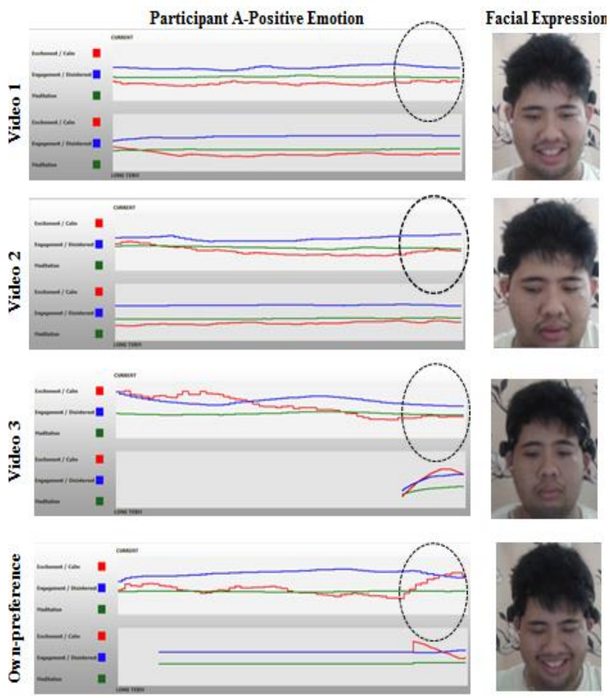

Figure 6: Brainwave graphs and facial expression of participant A while being induced with positive emotion videos.

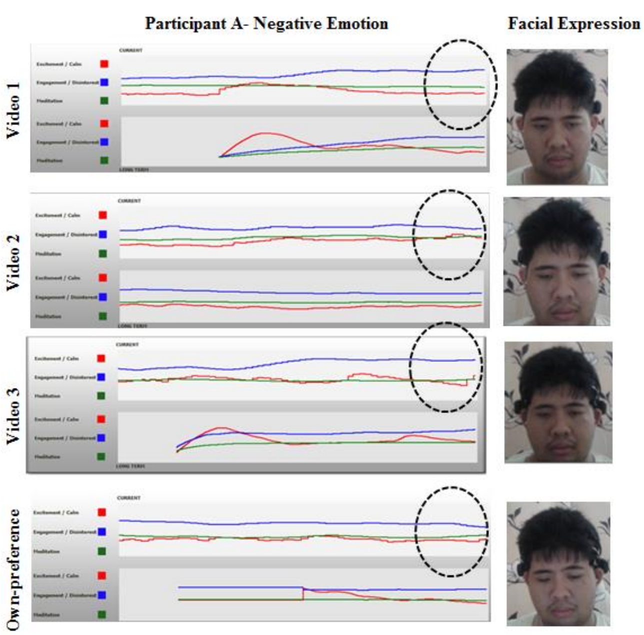

Figure 7: Brainwave graphs and facial expression of participant A while being induced with negative emotion videos.

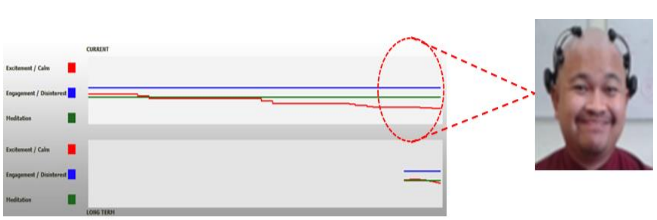

Figure 8: Participant C tries to mimic happy face.

IV. CONCLUSION

This paper is oriented for three purposes which are to figure out suitable method and design of experiment for obtaining spontaneous facial expression, to prove the reliability of the Emotiv Epoc to detect the brain signals according to the emotion as well as to collect genuine facial expression images as data base training. From this pilot study, we found out that emotion is best to be induced by using stimulus which is matched with the person's preference.

Next, it is deduced that Emotiv Epoc ability is proven to detect the authenticity of human emotion and their expression by showing changes in its graph. Moreover, even a person posed his or her expression, the brainwave would not align with as the posed expression. This happened because the graphs change solely based on what people think and feel. Therefore, by using Emotiv Epoc, it is possible to collect spontaneous facial expression. Besides, we also found that, even someone tries to mimic an emotion, the expression display has slight differences in their features intensity as compared to real expression.

As a conclusion, this preliminary study has achieved its aims. For future work, we plan to explain in details regarding the spontaneous database that we have obtained in our next paper. Later, the extracted facial expression will be used for completing the next phase in our study which is to extract facial features related to the emotions and carry out the emotion classification.

ACKNOWLEDGMENT

We would like to thank all people and UTeM who directly or indirectly provide support for this work.

REFERENCES

[1] Verma G.K. and Tiwary U.S "Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals". NeuroImage, 1022014. 2013. pp.162–172.

[2] K. Takahashi, "Remarks on emotion recognition from multi-modal bio-potential signals", IEEE International Conference on Industrial Technology, IEEE ICIT Vol 3. 2004. pp.95–100.

[3] S. Kumar, T. K. Das and R. H. Laskar, "Significance of Acoustic Features for designing an Emotion Classification System", International Conference on Electrical and Computer Engineering (ICECE), 2014. pp.128–131.

[4] C. Houwei, A. Savran, R. Verma and A. Nenkova, "Acoustic and lexical representations for affect prediction in spontaneous conversations", Computer Speech and Language, 29(1), 2015, pp.203–217.

[5] Ten Bosch, L., "Emotions, speech and the ASR framework. Speech Communication", 40, 2003, pp.213–225.

[6] Lv. Guoyun, S. Hu and Lu. Xipan, "Speech Emotion Recognition Based on Dynamic Models" International Conference on Audio, Language and Image Processing (ICALIP), 2014, pp.480–484.

[7] L. Diane, F. Kate and S. Scott., "Towards Emotion Prediction in Spoken Tutoring Dialogues" 2012, pp.52–54.

[8] D. J. Litman and K. Forbes-Riley, "Predicting student emotions in computer-human tutoring dialogues", Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics - ACL '04, 2004, p.351.

[9] J. Lin, A, C. Miao, Z. Shen, "A FCM based Approach for Emotion Prediction in Educational Game", 7th International Conference on Computing and Convergence Technology (ICCCT), 2012, pp. 980–986.

[10] S. Yecker, J. C. Borod, A. Brozgold, C. Martin, M. Alpert and J. Welkowitz, "Lateralization of Facial Emotional Expression in

Schizophrenic and Depressed Patients" J Neuropsychiatry Clin Neurosci, 1999, pp.370–379.
[11] Daniel. J. France, Richard G. Shiavi, S. Silverman, M. Silverman and D. Mitchell Wilkes, "Acoustical Properties of Speech as Indicators of Depression and Suicidal Risk", IEEE Transaction on Biomedical Engineering, Vol.47(7), 2000, pp.829–837.
[12] X. Huang, D. Chen, Y. Huang and X. Han., "Automatic Prediction of Trait Anxiety Degree Using Recognition Rates of Facial Emotions", Sixth International Conference on Advanced Computational Intelligence (ICACI), 2013, pp.272–275.
[13] K. Holewa and A. Nawrocka, "Emotiv EPOC neuroheadset in brain - computer interface", Proceedings of the 2014 15th International Carpathian Control Conference (ICCC), 2014, pp.149–152.
[14] J.Benas, (2014), "A Blind Father and His Daughter-Short Sad Story", Retrieved from https://www.youtube.com/watch?v=SV6Sw9CNWek>.
[15] Changing Batteries- The Saddest Story 3D Animation, (2013), Retrieved from https://www.youtube.com/watch?v=O_yVo3YofqQ>.
[16] Twin Babies Fight Over Pacifier, (2013), Retrieved from https://www.youtube.com/watch?v=NfZC44C6aCU>
[17] Emmerson-Mommy's Nose is Scary (2011), Retrieved from https://www.youtube.com/watch?v=N9oxmRT2YWw>.W.-K. Chen.