Emotion Recognition Using Video-Watching EEG Based on Late Positive Potential

Zhang Luyang*1, Yuuki Noro*2, Tomohiro Yoshikawa*3, and Takashi Furuhashi*4

*1*2*3*4 Dept. of Information and Communication Eng., Graduate School of Eng., Nagoya University
*1E-mail: zhang.luyang@j.mbox.nagoya-u.ac.jp
*2E-mail: noro@cmplx.cse.nagoya-u .ac.jp
*3E-mail: yoshikawa@cse.nagoya-u .ac.jp
*4E -mail: furuhashi@cse.nagoya-u .ac.jp

Abstract. With the development of brain–computer interactions, emotion recognition has become one of the most important fields of research in brain sciences. The EEG can clearly reflect the changes in the psychological and physiological state of the person. Therefore, it is important for researchers to evaluate emotion recognition by EEG signals. In previous studies, we considered casual sensing and performed emotion recognition by using a single electrode in the occipital region (Pz) the basis of an image dataset, used the International Affective Picture System, and collected the Late Positive Potential (LPP) and the EEG signal when a participant was stimulated by negative and positive emotions. By using the LPP data computed by the arithmetic mean of 10 sets of LPP data from a Pz electrode, we achieved an accuracy of 0.748 on the relevance ratio depression of neutral [1]. We then tried to extract the LPP from a video-based emotion recognition dataset by template matching. The DEAP dataset [2] was used and yielded an average accuracy of 80.3% and 82.5% with arousal and valence, respectively.

Keywords: Late Positive Potential, Event-Related Potential, Brain–Computer Interface, Emotion Recognition

1. INTRODUCTION

Emotion recognition was ignored by cognitive science researchers until the end of the twentieth century mainly because of the difficulties of data collection. With the advancement of technology, the research on Human Computer Interaction (HCI) has attracted significant attention in recent years. Emotion recognition is one of the research directions in HCI and plays an important role in HCI studies. By using emotion recognition, a number of applications has been created and applied in society, such as Face Targeting AD, which is a method that can recognize the emotion of users by camera and recommend advertising automatically. Kanjoya is a type of robot that uses sound signals to perform emotion recognition and change the tone of speaking during conversations. With the advent of portable electroencephalography (EEG) signal measuring instruments, it is now possible to collect and use brain signals on a daily basis.

At present, we mainly use electroencephalograph as acquisition equipment for collecting brain signals. EEGs record weak brain signals via electrodes placed on the scalp and then couple the signals to a differential amplifier via the electrode leads. Thereafter, the digitized signals would saved in a supporting system (e.g., a PC). Regarding the position of the scalp electrode, many placement methods have been proposed, such as the Montreal method, the Cohn method, and the Gibbs method. The most widely used method is the 10/20 system method, which was proposed by the International Federation of Clinical Neurophysiology. The electrode number is often set as 16, 32, 64 or 128. In recent EEG headsets, such as neurocam, few electrodes are used in collecting the brain signals.

Previous research shows that when the volunteers are stimulated by image signals, there may be a certain noticeable tendency in the occipital region. One of them is event-related-potential (ERP). For example, it has been reported that an ERP called late positive potential (LPP) is observed when feeling negative or positive emotions evoked by the images of International Affective Picture System (IAPS) [3].Therefore we focus on the possibility that this LPP is a characteristic component due to emotion. In previous studies, we used LPP data that was collected by a Pz electrode based on IAPS to predict the score of neutral by stepwise linear discriminant analysis (SWLDA) and Linear regression, we obtain an accuracy of 0.748.The EEG signal is a low S/N ratio signal mixed with lots of noise, such as EOG and EMG. To our knowledge, previous studies presented many methods for denoising, but removing the background noise thoroughly is still difficult. Thus, we let every volunteer watch the dataset for 10 times, and then computed the arithmetic mean of every 10 sets of data. As the result, the accuracy increased by 23.3 %.

However, in practical applications, we have to repeatedly show the same image to the participants to compute the arithmetic; this process may make the user feel bored, and the image may not be able to evoke emotions. Furthermore, even if the time of image stimulation is shortened in comparison to the conventional study, the test time is still too long to the participants; therefore, the condition of the participants may adversely affect the EEG signal.We tried to perform emotion recognition by using the Dataset for Emotion Analysis using EEG, Physiological, and Video Signals (DEAP) dataset, which was collected when the participants were watching the music videos. The use of video stimulation not only reduces the annoyance of the participant but also greatly reduces the experimental time because it is not necessary to show the images repeatedly to the participants.

We used only the Pz electrode to recognize the emotion by LPP and arithmetic mean.

2. DATASET

The DEAP is a bio-signal open database for emotion analysis. It consists of the EEG signals of 32 participants, as well as peripheral signs (eyes, mouth, shoulders, fingers, etc.) and presented MVs. The MV was shown to the participants for 1 minute to record the EEG, and there are 40 MVs in total. Each MV was selected by a preliminary experiment. The age range of the participants is 19 to 37 years old (with a mean of 26.9 years old). For 22 of the 32 participants, frontal face video was also recorded. EEG and peripheral signals were recorded at a sampling rate of 512 Hz, but were downsampled to 128Hz. The eye artifacts were then removed, and the high-pass filter was applied. For each participant, there is a file that contains two arrays with the shape of (40×40×8064) and (40×4).

Name	Array shape	Array shape
data	40 × 40 × 8064	video/trail × channel × data labels
labels	40 × 4	video/trail × label

There is a 3-second pretrial baseline before every signal, and the 40 channels contain 8 channels records (EOG, EMG, GSR, etc.). Thus, we cut the first 3-seconds of data and use the data array in the shape of (40 × 32 × 7680).

3. LPP

LPP is a type of ERP that would be induced by an image that stimulates emotions [4]. LPP is a feature component that appears in the low frequency band of approximately 1Hz-10 Hz, in which a positive potential change occurs after approximately 200 ms - 500 ms of image stimulation. Considering that LPP is a weak signal with a low S/N ratio, it is difficult to capture it with a single trial. Therefore, it is necessary to clarify LPP by adding and averaging a plurality of brain waves to improve the detection rate. In a previous study [1], the following two results were obtained by using images belonging to the three emotion categories of negative (unpleasant), neutral, positive (pleasant) among IAPS images.

3.1. The Region of LPP Measuring

The amplitude of the measured LPP increases as it moves from the forehead to the occipital region. The three brain waves in Figure 1 show the electroencephalogram measured from the top to the forehead (Fz electrode), the top (Cz electrode), and the occipital (Pz electrode). It can be confirmed that the amplitude increase of LPP at the Pz electrode is more significant than that at the Fz and Cz electrodes.

Additionally, Figure 2 shows a topo-map that indicates the relation of magnitude of the LPP amplitude seen from the overhead direction, in which red represents large amplitude and blue represents small amplitude. LPP is weak in the vicinity of the temporal region; thus we can see that the above tendency is mainly on the midline. To summarize, we should focus our attention on the brain waves measured in the occipital region on the midline to understand the LPP data. Emotion recognition can be performed only by the Pz electrode, which was normally installed on portable devices.

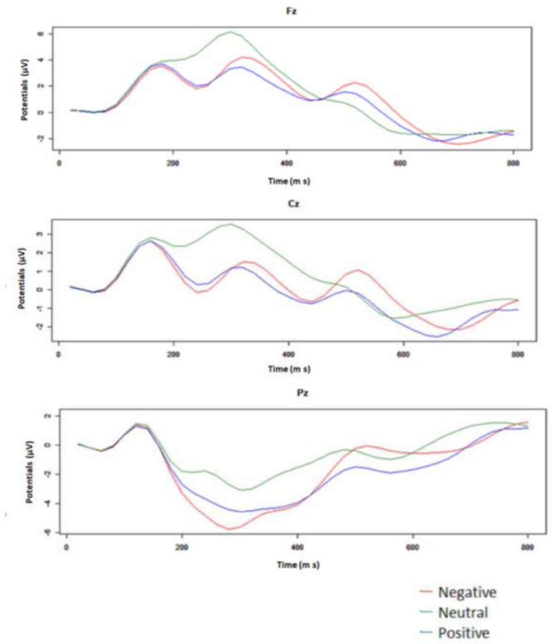

Fig.1 LPP wave collected by Fz, Cz and Pz

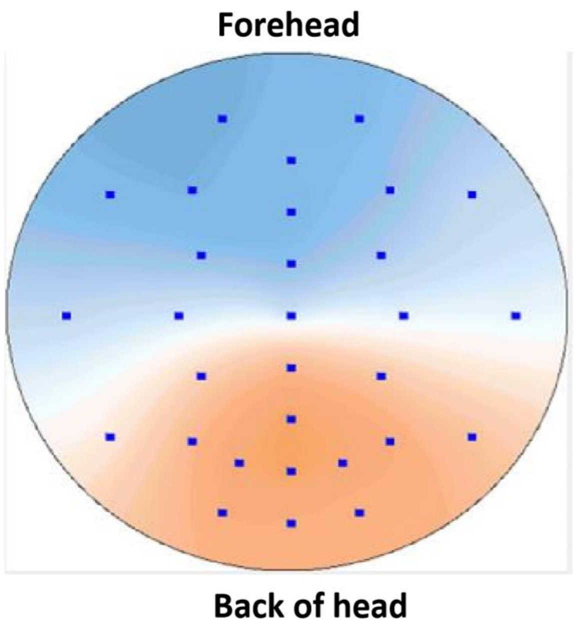

Fig.2 LPP amplitude topo-map

3.2. Neutral Emotion Recognition by LPP

By comparing the amplitudes of LPP when displaying negative and positive images, the amplitude of neutral is suppressed. In Figure 1, it is confirmed that the amplitude of neutral is suppressed by comparing the LPP measured by the Pz electrode of negative and positive. Therefore, it is possible to judge if a participant is in the emotional state of neutral.

4. RELATED WORKS

Koelstra et al. [5], who are the providers of DEAP dataset, computed the accuracy of emotion recognition by using a naive Bayes classifier based on multimedia content analysis [6].

A previous study, used the three teacher labels (arousal, valence, and liking) and computed the average power spectrum of the four frequency bands (4 – 7 Hz, 8 – 13 Hz, 14 – 29 Hz and 30 – 47 Hz) of all electrodes, and showed the trend difference by topo-maps.

Yoon and Chung [7] proposed a new methodology for emotion recognition from EEG signals. The DEAP dataset was used to verify this method. Fast Fourier transform analysis was used in feature extraction, and feature selection based on Pearson correlation coefficient was applied on the extracted features. They proposed a probabilistic classifier based on Bayes theorem and supervised learning by using a perceptron convergence algorithm.

5. PROPOSED METHOD

5.1. Description

The purpose of this study is to improve estimation accuracy by using the averaging method and LPP for video stimulation. Moving images are continuous images; thus, LPP is elicited when images that evoke emotions are presented. A section that is considered an LPP section is extracted by using template matching from brain waves for 60 seconds, and the section is analyzed.

We used 10 sets of LPPs data in total, including data from 10 subjects data obtained in the previous study [1] to construct the template of LPP. LPP is an ERP and is a feature component that is consistently obtained regardless of the type of subject even though it has individual differences. Therefore, it seems unlikely that it will be a completely different tendency in comparison with another subject group. Figure 3 shows the waveform of LPP that was used as a template and the standard deviation at each time divided by 10. This means that the average and standard deviations from the stimulus presentation of 8000 samples of 10 subjects were measured at the Pz electrode at the time of displaying the negative and positive images in the used LPP to 1000 ms. By using this waveform as a template, the brain waves in the Pz electrode of DEAP that were included in the range obtained by integrating the constant k with the standard deviation were extracted as LPP. Thereafter, for each 60-second brain wave, we performed table matching in each window with a length of 1000ms and the stride of 31.25 ms, as it shown in fig.4.

During table matching, we set the beginning point of the windows as 0 μV. Furthermore, k was set by each participant with an increment of 0.1 and was determined by the smallest k that can satisfy the minimum extraction number of an LPP was larger than 5 among 40 MVs.

In this paper, stepwise linear discriminant analysis (SWLDA) method was used for classification. At present, SWLDA has been widely used in P300 based BCI and has achieved good classification results [9]. The emotion classification model is constructed by using the average value of LPP waveforms extracted before as the feature components and using SWLDA to classify them.

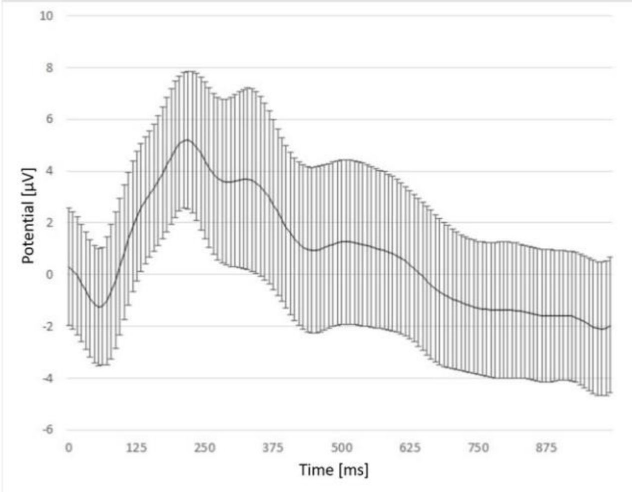

Fig.3 LPP Waveform

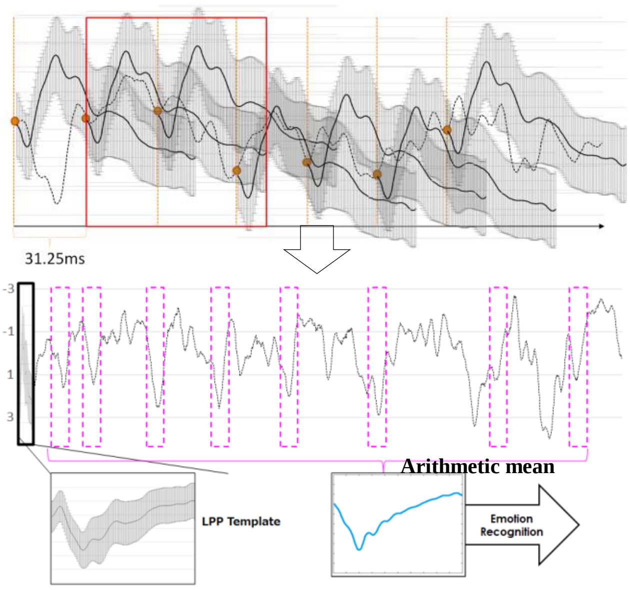

Fig.4 Template matching and LPP extraction

5.2. Experimental

First, to remove the AC wave components, we used a 50 Hz notch filter and downsampled the sampling frequency into 128 Hz. We extracted the LPP by table matching as described in Section 5.1. We then used a different k table match on the denoised data by computing the arithmetic mean of 32 participants for each video and compared the number of LPPs extracted from the data. Table 2 shows a part of the results.

Table. 2 The number of LPP.

K	average	Minimum number of extractions
0.11	1.2	0
0.12	4.7	0
0.13	14.0	7
0.14	32.9	20
0.15	68.7	42

The table shows that it is possible to acquire the feature of positive potentials when $k > 0.13$. We then compared the arousal and valence label score of each MV with a score of 5.0. If the score > 5, the score is set to "1"; otherwise the score is set to "0". We identified the model for each participant by the label and the arithmetic mean of the LPP extracted from the dataset by k. The accuracy is then computed using SWLDA and 10 split cross-examination. For comparison, we also computed the accuracy by using a previous research method [7] and the proposed method. As a result, we achieved higher accuracy on the basis of LPP data. The stepwisefit function of Matlab 2014a was used in this research with the argument penter = premove = 0.05. The result is shown in figure 5 and table.3. The accuracy of both arousal and valence were higher than that of the estimated model of the conventional study. We showed that LPP extraction by template matching and its averaging emotion estimation method are more effective than conventional methods. In future works, we plan to apply and consider methods that reduce noise without applying the averaging method (e.g., signal source separation).

Table.3 The average of the accuracy and standard deviation

	Arousal	Valence
LPP	0.736(0.101)	0.816(0.043)
Previous study	0.536	0.702

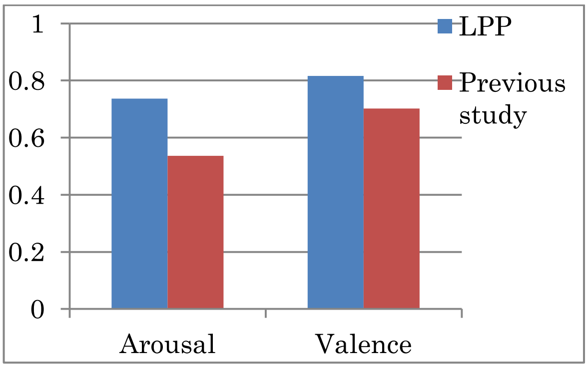

Fig.5 Accuracy comparison

6. CONCLUSION

This paper presents a discussion on the possibility of extracting LPP by using one electrode to recognize emotion on the basis of a video-watching EEG dataset. The results show that LPP may recognize more expressive features in emotion recognition than previous studies. It is desirable to perform template matching by using the LPPs of the same subject, and it is necessary to repeat the EEG measurements. Furthermore, template matching is applied to brain wave data that contains a large amount of noise; thus, brain waves that are not LPP originally were extracted. Applying techniques to eliminate or reduce this noise is expected to further improve the accuracy of emotion estimation.

7. ACKNOWLEDGMENTS

This work was supported by the Center of Innovation Program of Japan Science and Technology Agency.

[1] Y. Noro, T. Yoshikawa, T. Furuhashi, "A Study on EEG-based Emotion Recognition using Averaging Process," 2017 9th International Conference on Soft Computing and Intelligent Systems, 2017

[2] S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, "DEAP: A Database for Emotion Analysis using Physiological Signals (PDF)," IEEE Transaction on Affective Computing, Special Issue on Naturalistic Affect Resources for System Building and Evaluation, in press K. Elissa, "Title of paper if known" unpublished.

[3] Peter J Lang, Margaret M Bradley, and Bruce N Cuthbert. International affective picture system (IAPS): Affective ratings of pictures and instruction manual, 2008.

[4] Picton, Terence W. "The P300 wave of the human event-related potential." Journal of clinical neurophysiology 9.4 (1992): 456-479.

[5] Sander Koelstra, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras. Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, 3(1):18–31, 2012.

[6] Alan Hanjalic and Li-Qun Xu. Affective video content representation and modeling. IEEE Transactions on multimedia, 7(1):143–154, 2005.

[7] Hyun Joong Yoon and Seong Youb Chung. Eeg-based emotion estimation using bayesian weighted-log-posterior function and perceptron convergence algorithm. Computers in biology and medicine, 43(12):2230– 2237, 2013

[8] Peter J Lang, Margaret M Bradley, and Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2006). A Comparison of Classification Techniques for the P300 Speller. Journal of Neural Engineering, 3(4), 299.

[9] MCCANE L M, HECKMAN S M, MCFARLAND D J, et al. P300-based brain-computer interface (BCI) event related potentials (ERPs): People with amyotrophic lateral sclerosis (ALS) vs . age-matched controls [J]. Clinical Neurophysiology, 2015, 126(11): 2124-2131.