

Features of Event-related Potentials Used to Recognize Clusters of Facial Expressions

Masahiro Yasuda¹, Zhang Dou² and Minoru Nakayama²

¹*Mechanical and Control Engineering, Tokyo Institute of Technology, O-okayama, Meguro-ku, Tokyo, 152-8552 Japan*

²*Human System Science, Tokyo Institute of Technology, O-okayama, Meguro-ku, Tokyo, 152-8552 Japan*

Keywords: Human Emotion, Facial Expression, EEG, ERP, Chronological Analysis, Prediction.

Abstract: To assess human emotion using electroencephalograms (EEGs), the relationship between emotional impressions of images of facial expressions and features of Event Related Potentials (ERPs) recorded using three electrodes was analyzed. First, two clusters of emotional impressions were extracted using two-dimensional responses of the Affect Grid scale. Second, features of ERPs in response to the two clusters were examined. Time slots where amplitude differences in ERP appeared were measured, and differences in the frequency power of ERP were also extracted for each electrode. To evaluate these features, prediction performance for the two clusters was examined using discriminant analysis of the features. Also, the dependency of some band pass filters was measured.

1 INTRODUCTION

In order to develop good Human-Computer interfaces and create good communications systems, such as human-robot or human-human, the establishment of a technique for assessing human emotion is necessary. Various biosignals have been used to try to detect human emotion (Lin et al., 2010). Electroencephalograms (EEG) are a type of bio-signal sometimes used to detect human emotion, the feature extraction and signal processing procedures of which have been previously discussed (Guitton, 2010; Petrantonakis and Hadjileontiadis, 2010). In those experiments, categories of emotion were based on Ekman's classification system (Ekman and Friesen, 1975), and also on visual stimuli which can evoke the viewer's impression of the emotion they are viewing. A simple stimulus for presenting emotion is a set of facial expressions which stimulate the viewer's impression of the emotion they are viewing. Even during the recognition of facial emotions, the transformability of emotions and individual differences should be considered, however. The relationship between certain facial expressions and the viewer's impression of the emotion they are viewing is sometimes analyzed (Huang et al., 2009). Therefore, biosignal reactions such as EEGs should be analyzed in response to the viewer's impression of the emotion they are viewing. To extract specific physiological components of stimuli, Event

Related Potentials (ERPs) are often used to analyze EEGs. ERP waveforms are averaged waveforms of EEGs which respond to stimulus. ERPs are often used for chronological analysis in psychological and clinical studies (Nittono, 2005; Rugg, 1997). Also, viewer's chronological reactions to emotional stimuli can be analyzed using ERP waveforms. Analysis of these can contribute to the extraction of good features used to detect viewer's emotional responses.

To examine the relationship between the observer's impressions of emotional face images and ERP responses when viewing this stimuli, significant feature information from the ERP waveforms was extracted in an experiment. Therefore, this paper addresses following topics:

1. The groups of facial emotions which are based on viewer's impressions were extracted using subjective evaluation scores.
2. The ERP waveforms which are based on the emotion groups were compared and the differences were extracted.
3. To evaluate the significance of feature differences in ERP waveforms, the performance of discriminant analysis was evaluated. Also, the dependency of some band pass filters was measured.

For these purposes, the following experiment was conducted.

2 EXPERIMENTAL METHOD

2.1 Stimulus

The images of facial emotions in this experiment were prepared using the Japanese and Caucasian Facial Expression of Emotion (JACFEE) collection (Matsumoto and Ekman, 1988). This collection consists of 56 color photographs of 56 different individuals who illustrate one of the seven different emotions: Anger, Contempt, Disgust, Fear, Happiness, Sadness and Surprise. The photos are of equal numbers of Japanese and Caucasian models: 14 Caucasian males, 14 Caucasian females, 14 Japanese males, 14 Japanese females. The validity of the facial emotions has been confirmed, but some confusion in recognizing some of the expressions has also been reported (Huang et al., 2009).

The subjects (viewers) who participated in this experiment were 6 university students from 19 to 23 years old. Their visual acuity was sufficient to observe the stimuli. The experimental content was explained to all participants in advance, and informed consent was then obtained.

2.2 Procedure

Facial photos as visual stimulus were presented sequentially. The experimental sequence is illustrated in Figure 1. First, a fixation point (+) was shown in the centre of the screen, and a stimulus appeared after the screen had been blank for 3-4 seconds. A fixation point was displayed to attract eye fixation, but it disappeared before the target stimulus was shown, in order to trigger stimulus onset in the manner of a conventional visual perception experiment (Kirchner and Thorpe, 2006). The duration of stimulus display was 3 seconds during which the PC sent a trigger signal to another recording device. The display sequence was controlled using Psychtoolbox (Brainard, 1997). A set of sequences consisted of 56 photos, and the total duration was 6 minutes. Three trials were conducted in which different sets are shown to each subject, followed by short breaks.

The EEGs were recorded from 3 scalp electrodes positioned in the Frontal (Fz), Central (Cz) and Occipital (Oz) areas, according to the international 10-20 system. The EEG potentials were measured using a bio-amplifier (ADInstruments: PowerLab4/30, ML13). The scalp electrodes were referenced to a base measurement at the subject's ear lobes. A ground electrode was placed on the forehead. The following sampling conditions were used to record signals on a PC, sampling rate: 400Hz, low pass filter: 30Hz, time

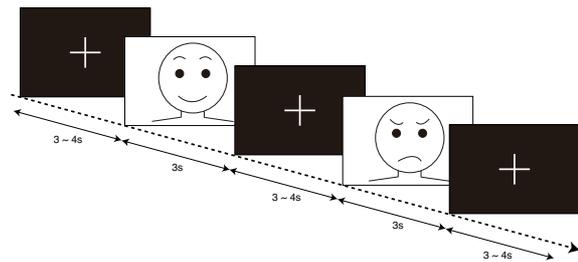


Figure 1: Diagram of showing stimuli.

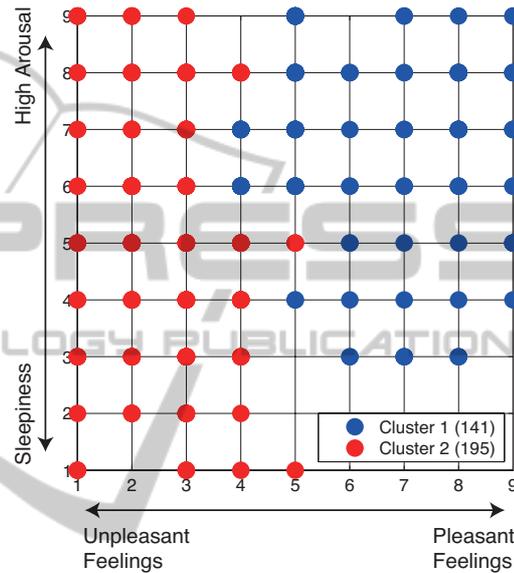


Figure 2: Results of cluster analysis for viewer's responses using the Affect Grid.

constant for a high pass filter: 0.3sec. To detect blinks as an artifact source, in addition to the three potentials the vertical component of an electro-oculograph (EOG) was measured synchronously.

Additionally, five sets of band pass filters were used in this analysis to reduce the artifacts of lower frequencies in EEGs. The filter was applied to signals during off-line processing using LabChart (ADInstruments), which is a zero-phase-lag Finite Impulse Response (FIR) filter. The ranges of band passes were controlled using the following frequency bandwidths : 0.5 ~ 30Hz, 1.0 ~ 30Hz, 1.5 ~ 30Hz, 2.0 ~ 30Hz, 2.5 ~ 30Hz.

2.3 Subjective Evaluation of Facial Emotions

In order to evaluate the viewer's impression of the emotion they are viewing of the stimulus photos, all subjects were asked to rate the emotions using a designated scale. The scale is known as an "Affect Grid",

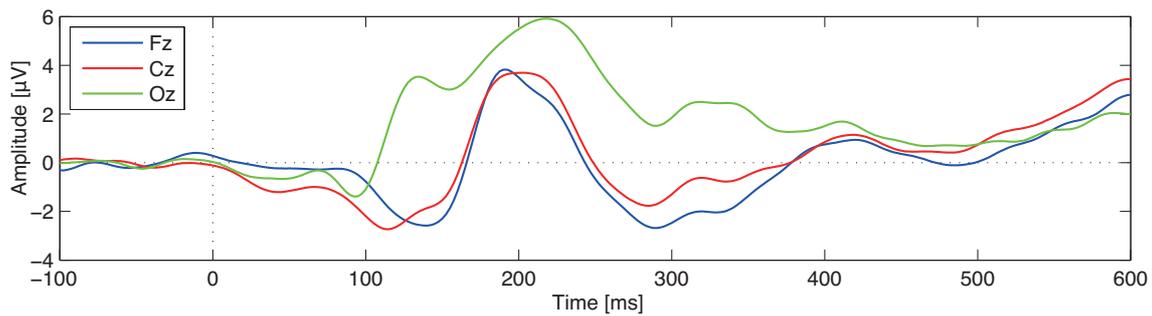


Figure 3: Three ERPs for Electrodes Fz, Cz and Oz (Band Pass Filter: 2.0 ~ 30Hz).

and consists of a two dimensional 9 point scales with “Pleasant - Unpleasant Feelings” and “High Arousal - Sleepiness” (Russell et al., 1989), as shown in Figure 2. The validity of this scale has often been used to evaluate facial emotions (Takehara and Suzuki, 2001; Shibui and Shigemasu, 2005).

All 56 photos were rated by each subject using the scale, after observing three trial sets of images.

3 RESULTS

3.1 Responses using the Affect Grid

All ratings for the facial image photos using Affect grid are summarized in Figure 2. The horizontal axis indicates “Pleasant - Unpleasant Feelings”, and the vertical axis indicates “High arousal - Sleepiness”. In regards to Figure 2, the ratings cover most of the area of the two-dimensional scale in response to the photos of 7 facial emotions. The responses for two-dimensional scales were given using integer grading, and many of the responses overlap others in Figure 2. In comparing the rating distribution between participants across two-dimensions, some individual differences were observed (Yasuda et al., 2014). Though facial codes are universal emotions used in Ekman’s study, there are some individual differences in the ability to recognize them. Therefore, the subject’s responses were not identical for each photo.

To classify the responses based on their two-dimensional ratings, cluster analysis using the weighted pair group method and arithmetic averages (WPGMA), which is a distance metric used to facilitate comparison with each other, was conducted for all 336 responses (56 photos \times 6 subjects). As a result, two clusters were extracted, and are presented as two colors in Figure 2. Two clusters are distributed in the left and right regions of the horizontal axis, which shows “Pleasant - Unpleasant Feelings”. Therefore, they can be called “Pleasant” and “Unpleasant” clusters. The ratio of the number of clusters depends on

each subject, and the average ratio for “Pleasant” is 0.42.

3.2 Event Related Potentials

3.2.1 Chronological Analysis of the Two Clusters

To measure the differences in EEG waveforms between “Pleasant” and “Unpleasant” clusters, the two were compared chronologically. Here, event related potentials (ERPs) are calculated to emphasize the waveforms in response to the two classes of photos of facial emotions.

EEG waveforms were extracted between -100 and 600 milliseconds before stimulus onset, and the average voltage in the 100 milliseconds before stimulus onset was set as a baseline. Trials which contained artifacts such as blinks or amplitudes over $\pm 100\mu V$ were excluded in advance, and averaged potentials for the two clusters were then calculated.

First, the ground averages of ERPs are summarized to compare the waveforms of the three electrodes in Figure 3 chronologically. The deviations start around 100 milliseconds for all ERPs, and this phenomenon has been observed in previous studies. The first negative peaks were observed around 100 milliseconds in the order of Oz, Cz and then Fz. This may suggest that visual evoked signals spread from the Occipital area (low level vision) to the Frontal area (high level vision). A broad range of waveforms is observed for Oz while voltage changes for Cz and Fz appear at around 200 milliseconds. These responses may reflect to the progress of visual information processing.

Second, the ERP waveforms between the two clusters on each electrode are compared in Figure 4. The solid line represents the “Pleasant” cluster, and the gray line represents the “Unpleasant” cluster. There are some differences in ERPs between the two clusters. To identify the time zone where there are significant differences between the two means of ERPs across the two clusters, pair-wise t tests were

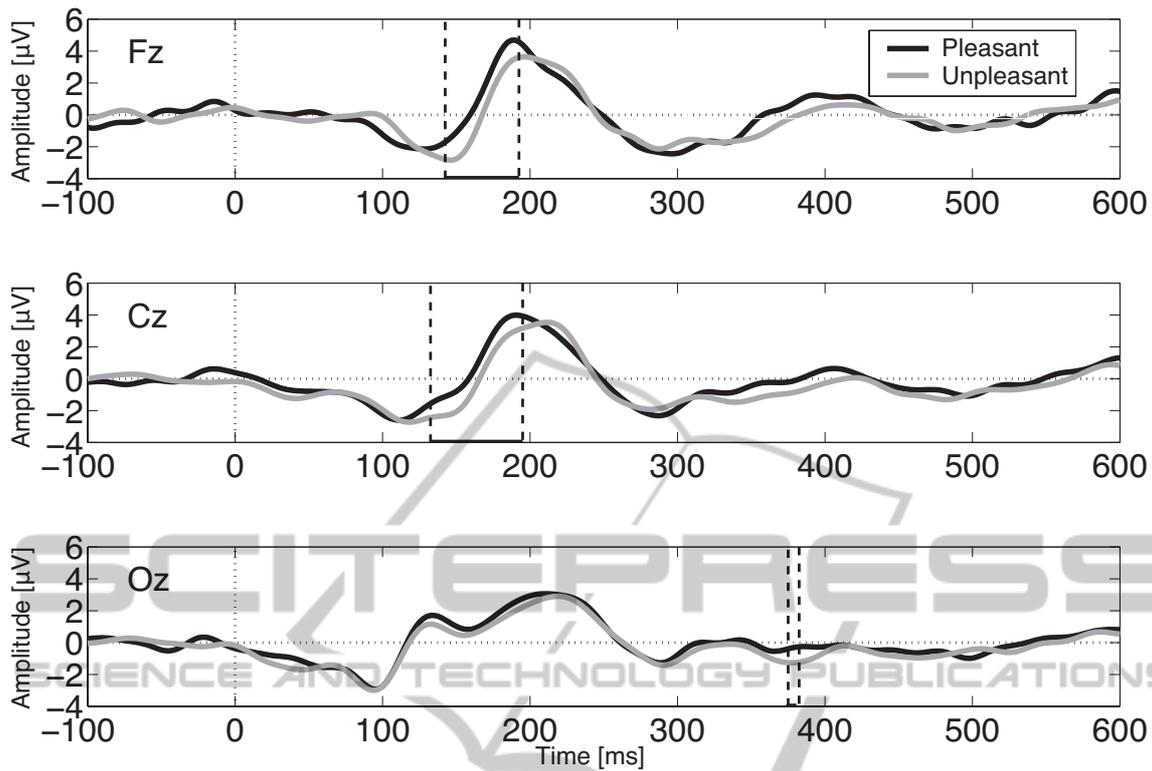


Figure 4: Comparison of ERPs between two clusters of Electrodes Fz, Cz and Oz (Band Pass Filter: 2.0 ~ 30Hz), with a dotted-line box indicating the time zone where there are significant differences between two ERPs ($p < 0.05$).

conducted. In regards to previous studies (Thorpe et al., 1996; VanRullen and Thorpe, 2001), the time zone was identified where 15 consecutive t test values were below the $p < 0.05$ level. In this paper, a 5% level of significance was employed, and df was 34 (2 clusters \times 3 sets \times 6 subjects - 2). The sampling slot mentioned above is 2.5 milliseconds.

The significant time slots can be extracted, and they are illustrated using dotted line boxes in Figure 4 as follows, Fz: 142.5~192.5 milliseconds, Cz: 132.5~195.0 milliseconds. Both are time slots occurring after the negative peaks at around 100 milliseconds and before the positive peaks at around 200 milliseconds. Those time slots are independent of the type of band pass filter. Therefore, ERPs in those time slots may contain some features in response to the two clusters of emotions. At the Oz electrode, the significant time slot is too short and too late, so it should be ignored.

3.2.2 Effectiveness of Band Pass Filters

A band pass filter is used in EEG measurements to detect a distinct signal. To determine the appropriate filter band to extract features of emotional responses, the variance of these signals were evaluated using

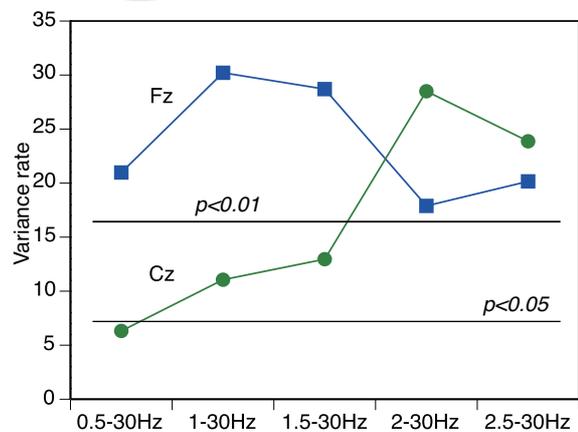


Figure 5: The rate of variance between two conditions.

analysis of variance (ANOVA). To maximize the variance between the two clusters, the appropriate band pass filter was evaluated using the following procedure. In addition to the randomized factor of the viewers, two factors, namely the two clusters and the trial sets (2 \times 3), influenced the deviations of the potentials in this experiment. The ratio of variance between the two clusters (F values) was calculated for the various filter conditions.

F values are summarized in Figure 5. The hor-

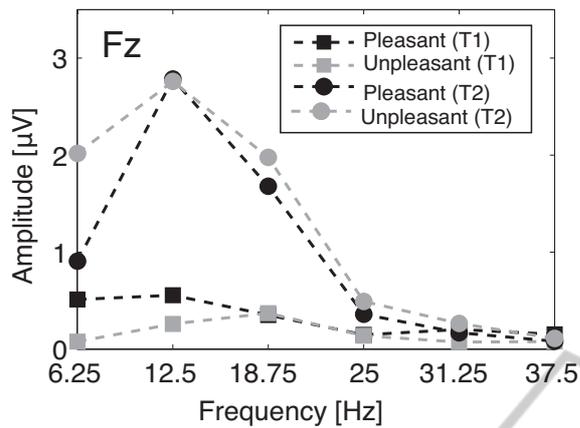


Figure 6: Comparison of frequency spectrum for Electrode Fz across two conditions and time slots (Band Pass Filter: 2.0~30Hz).

horizontal axis represents filter conditions, the vertical axis represents F values. Two levels of significance in F values are indicated in Figure 5. The results show that all conditions for Fz always produce a significant variance ratio, but Electrode Cz needs filters above 2Hz. Therefore, the band pass filter should be set over 2Hz to detect the differences in ERPs between the two clusters.

Additionally, the factor of the trial set is not significant and does not influence other factors in this analysis.

3.2.3 Frequency Analysis of EEG

In the above section, amplitudes of ERP between the two clusters are compared. The frequency components is another well known feature of EEG/ERP signals, and the difference in these components should also be analyzed. Regarding the chronological analysis, typical ERP responses can be observed in the 200 milliseconds after stimuli onset in comparison with the baseline, which is before onset. To extract some features of ERP responses, simple frequency analysis was employed to obtain the frequency factors. Regarding measuring restrictions, the features were created using the following procedure. Frequency analysis was applied to two time slots, T1: -100~60 milliseconds and T2: 60~220 milliseconds, both were 160 milliseconds span (64 sampling points) whose length was based on a power of 2 using FFT algorithm. Hanning window was applied to FFT analysis. Regarding the analytical conditions, the frequency resolution is 6.25Hz.

The amplitudes in the frequency domain for Fz are summarized in Figure 6. The horizontal axis represents the frequency components which are based on the frequency resolution in this analysis. The vertical

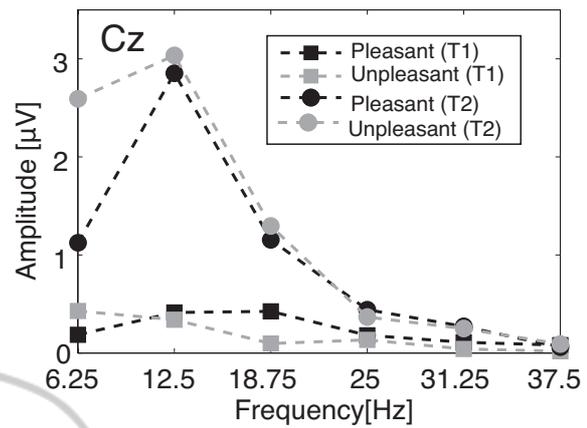


Figure 7: Comparison of frequency spectrum for Electrode Cz across two conditions and time slots (Band Pass Filter: 2.0~30Hz).

axis represents the amplitude of the components. The amplitudes in the frequency domain for Cz are summarized in Figure 7, and shown in the same format as Figure 6. The amplitudes of slot T2 are larger than the ones for slot T1, and the amplitudes of the first four components in T2 account for over 90% of the total. The differences in the amplitudes between the two clusters for both Fz and Cz at 6.25Hz are remarkable.

3.3 Estimation of Emotion Clusters using Features of ERP

There are some features of ERPs which present two emotion clusters, as mentioned in the above sections. To measure the significance of these features, discriminant analysis to predict the emotion clusters using the features was conducted. A Support Vector Machine (SVM) was applied to this prediction, and the subject leave-one-out procedure was employed to evaluate the estimation accuracy. The features follow 6 metrics as input data: amplitudes of ERP for each significant time slot for Fz and Cz, and amplitude differences in the lower four frequency components between time slots T1 and T2. The actual numbers of dimensions for features are 18 amplitudes of Fz, 25 amplitudes of Cz, and 4 frequency components of electrodes Fz and Cz, as mentioned in subsection 3.2.3. The number of dimensions is summed up when both Fz and Cz are employed. The feature set consists of 36 pieces of data (2 emotion clusters \times 3 set of trial \times 6 subjects). There are 6 prediction conditions: ERP amplitudes (Fz, Cz, and both) and Frequency components (Fz, Cz, and both).

The levels of accuracy are compared between conditions by employing five band pass filters mentioned

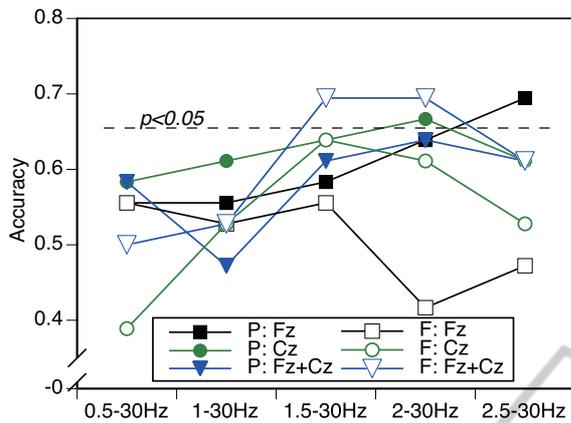


Figure 8: Accuracy of prediction for emotion clusters using features of ERPs.

above in Section 2.2. They are summarized in Figure 8. The horizontal axis represents the five band pass filters, and the vertical axis represents the accuracy. This accuracy gradually increases as the band pass filter conditions vary toward lower frequencies (0.5~2.5Hz). To evaluate the significance of the prediction, a χ^2 test was conducted. The level of significance in Figure 8 is indicated using a dotted line.

As a result, the level of accuracy of the four conditions is significant. Combinations of frequency components for Fz and Cz are significant for the level of accuracy while certain ERP amplitudes for Fz and Cz are effective. From these accuracy results, the lowest frequency of the band pass filter should be set at around 2.0Hz, as this will also maximize the variance in ERP amplitudes between conditions.

4 DISCUSSION

To evaluate viewer's emotion, a detectable class of emotions should be specified. In regards to the results of cluster analysis of responses using an Affect Grid, two clusters such as "Pleasant" and "Unpleasant" feelings can be extracted. When factors concerning the photo models were omitted, the same two clusters were also extracted. Though the classes of facial emotion are often discussed, typical emotions can be easier to detect. Therefore, the structure of the two clusters is an important scale for emotional impressions. Also, this classification can be applied to impressions for any types of images, such as "Pleasant" and "Unpleasant" images. The relationship between the two clusters of viewed facial emotions and the features of ERPs has been confirmed. The impression of the image can be estimated to be either of the two clusters using just the features of ERPs.

Regarding chronological analysis of ERP waveforms, significant differences in Fz and Cz ERPs can be detected between the two clusters, but there is no difference at Oz. The emotional recognition is one of high level processing, and the differences appear on potentials at the mid and frontal areas. Therefore, more detailed information should be collected from these areas of the scalp. Additionally, there is no significant difference in features of ERPs across the three trials. The validity of the possibility of detection has been confirmed. To detect this activity, the use of a selection of band pass filters directly affects feature extraction. Analysis of the amplitude and frequency components of ERPs confirms that the lower frequency of the band pass filter should be set at around 2.0Hz.

The possibility of predicting viewer's impressions and dividing them into two clusters of facial emotion has been examined. This means that viewer's subjective impressions of an emotion can be estimated using their EEG information while viewing occurs. To improve prediction performance, appropriate feature extraction and an appropriate combination of procedures should be determined. Also, the possibility of predicting the viewer's emotional condition will be confirmed using photos which exclude those which display facial emotions.

5 CONCLUSION

In order to evaluate the viewer's impression of the emotion they are viewing, the relationship between emotion impressions of images of facial expressions and features of Event Related Potentials was analyzed. Following points were extracted during the analysis.

1. Subjective impressions of the photos of facial images were measured using the Affect Grid, and two clusters known as "Pleasant" and "Unpleasant" clusters were extracted.
2. The time slots showed significant differences in ERP waveforms between the two clusters at the Frontal (Fz) and Central (Cz) electrodes, except at the Occipital electrode (Oz).
3. In comparing the frequency power of the two time slots, -100~60 milliseconds (T1) and 60~220 milliseconds (T2), the amplitudes of slot T2 are larger. The differences in the amplitudes of slot T2 between the two clusters for both Fz and Cz at 6.25Hz are remarkable.
4. Five band pass filters were used in this analysis. The results of the analysis confirms that the lower

frequency of the band pass filter should be set at around 2.0Hz.

Discriminant analysis was conducted using the features of ERP, in order to evaluate the significance of these, and some significant accuracy was obtained. To generate more significant features to measure the viewer's state of emotion while they are viewing images of facial expressions, some additional biosignals such as eye movements using EOGs should be considered. They will be subject of our further study.

ACKNOWLEDGEMENT

This research was partially supported by the Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Scientific Research (B-26282046:2014-2016).

REFERENCES

- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10:433–436.
- Ekman, P. and Friesen, W. V. (1975). *Unmasking the Face: A guide to recognizing emotions from facial clues*. Prentice-Hall, New Jersey, USA.
- Guitton, C. (2010). Emotions estimation from eeg recordings. Master's thesis, Imperial College, Londo, UK. http://www.academia.edu/1203149/Emotions_Estimation_from_EEG_Recordings.
- Huang, J., Fan, J., He, W., Yu, S., Yeow, G., Sun, G., Shen, M., Chen, W., and Wang, W. (2009). Could intensity ratings of Matsumoto and Ekman's JACFEE pictures delineate basic emotions? A principal component analysis in Chinese university students. *Personality and Individual Differences*, 46:331–335.
- Kirchner, H. and Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, 46:1762–1776.
- Lin, Y.-P., Wang, C.-H., Jung, T.-P., Wu, T.-L., and Jeng, S.-K. (2010). EEG-Based emotion recognition in music listening. *IEEE Transaction on biomedical Engineering*, 57(7):1798–1806.
- Matsumoto, D. and Ekman, P. (1988). Japanese and Caucasian facial expression of emotion (JACFEE) and neutral faces (JACNeuF). San Fransisco State University, San Francisco, CA, USA.
- Nitto, H. (2005). *Event-related potential guidebook for psychological research*. Kitaohji-shobo, Kyoto, Japan.
- Petrantonakis, P. C. and Hadjileontiadis, L. J. (2010). Emotion recognition from EEG using higher order crossings. *IEEE Transaction on information technology in biomedicine*, 14(2):186–197.
- Rugg, M. D. (1997). *Cognitive Neuroscience*. Psychology Press, UK.
- Russell, J. A., Weiss, A., and Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, 57(3):493–502.
- Shibui, S. and Shigemasu, K. (2005). A model of two-dimensional placement of the facial expressions of emotion. *The Japanese Journal of Psychology*, 76(2):113–121.
- Takehara, T. and Suzuki, N. (2001). Robustness of the two-dimensional structure of recognition of facial expression: evidence under different intensities of emotionality. *Perception and Motor Skills*, 93:739–753.
- Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381:520–522.
- VanRullen, R. and Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, 13(4):454–461.
- Yasuda, M., Dou, Z., and Nakayama, M. (2014). EEG features of recognizing facial expressions. Technical Report HCS2014-41, IEICE Technical report.