Feature Characteristics of ERPs and Eye Movements in response to Facial Expressions

Minoru Nakayama*, Masahiro Yasuda

Information and Communication Engineering, Tokyo Institute of Technology, Ookayama, Meguro, Tokyo, Japan

Abstract

The features of EEGs and eye movements were extracted to identify viewer’s emotional responses during the viewing of photos of facial expressions in order to understand the physiological reaction that occurs during the perception progress for emotions. Both EEGs and eye movements were measured using electro-oculograms (EOGs). Facial expressions from a photo database were classified into two groups according to the viewer's subjective evaluation of whether the facial images where “Pleasant” or “Unpleasant”. The group that the photos of facial expressions belonged to was predicted using the extracted features, and the prediction performance was analysed. A correlation analysis of the frequency powers extracted from EEGs and eye movements was also conducted, and the differences in relationships between the emotional categories was discussed. The results provide evidence of the chronological process during the perception of visual emotion and of the mutual EEGs and eye movement activity that these reactions produce.

Received on 21 September 2018; accepted on 29 October 2018; published on 16 January 2019

Keywords: ERPs, Eye Movements, EOG, Facial Expressions

Copyright © 2019 Minoru Nakayama and Masahiro Yasuda, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/eai.18-6-2018.156320

1. Introduction

Emotional impression is a kind of subjective assessment of an object, and the response to the impression creates a useful technique for the detection of emotional states, which are key to understanding human activity. When a response occurs, bio-signals are measured which can be used to evaluate the range of human emotional reactions [1]. Though emotional categorisations may deviate in certain cases, the categorisation of facial impressions has been relatively well studied and validated [2, 3]. Also, facial expressions affect a viewer's emotional state as a function of human communication. This phenomenon is recognised as emotional synchronisation [4], which is defined as the display of images of certain facial expressions that stimulate viewers to exhibit specific emotions. During the process of recognising facial expressions, our visual perception process works well automatically. The activity may be observed during a biological reaction, using an electroencephalogram (EEG) and by measuring eye movements. When analysed, the results may provide evidence that the information process used to recognise facial emotions can be applied to the recognition of human behaviour which reflects passion.

To examine the above hypotheses, both EEGs and eye movements were observed simultaneously during the presentation of images of facial expressions. The chronological relationships between them, such as the following, were analysed.

1. The features of EEGs and eye movements in response to the emotions the viewer recognised were extracted using analysis of time and frequency domain, such as differences in ERP or frequency spectrum powers of eye movements.

2. The contributions of features extracted from the prediction of viewed facial expressions were examined using prediction performance of the groups of facial impressions which were classified using the behaviour of the viewers.

3. The relationships between EEGs and eye movements, in particular the temporal changes in these relationships, were measured to observe the eye movement trigger of facial expressions.

*Corresponding author. Email: nakayama@ict.e.titech.ac.jp
The following sections will first briefly summarise topical papers related to the subject, and then present an experiment which observes viewer’s subjective evaluations of pictures of facial expressions and the biological responses, such as EEG and eye movements, that result. Following Section 3, the detailed responses and behaviours of both EEGs and eye movements are analysed, and a visual perception mechanism based on the observed data is discussed in Section 4. Finally, the extracted results are summarised in Section 5.

2. Related works

The recognition of facial expression has been studied as a function of human visual perception. Scan-path patterns of eye movements during the viewing of facial expressions were compared in order to understand their specific types of behaviour for perception of facial emotion [5], and the progress of information processing was discussed. Eye movements during the viewing of facial expressions were analysed in a time course (0-560 ms), and the scan-paths and the latency of eye movements, such as saccadic reaction times, were discussed [6]. Biological responses to facial expressions such as eye movements and event related potentials (ERPs), which are the accumulation of EEGs, have also been studied using time course analysis of the recognition process [7]. In regards to the study of ERPs which were extracted from chronological reactions (120-180 ms) during the recognition of facial expressions [8], the ERP differences between the two groups of facial expressions were examined [9]. These previous studies suggest that both eye movements and EEGs may be indices which reflect the viewer’s emotional state. The relationships between responses of EEGs and eye movements may provide a process for processing perceptual information and thus provide the possibility of predicting emotional states using this information. In addition, these results show that ERP responses to facial expressions occur earlier than the reactions to eye movements [6, 8], and that some specific areas of brain activation may trigger eye movement [9].

3. Experimental Method

3.1. Stimulus

Images of facial expressions employed as visual stimuli were prepared using the Japanese and Caucasian Facial Expression of Emotion (JACFEE) collection [2]. This collection consists of 56 colour photographs of 56 different individuals who illustrate one of the seven different emotions: Anger, Contempt, Disgust, Fear, Happiness, Sadness and Surprise.

The experimental sequence is illustrated in Figure 1. Each stimulus was displayed for 3 seconds following a blank image used to produce eye fixation. The subjects were not asked to respond without viewing a facial image. Therefore, the facial images were displayed as large as possible on a LCD monitor, so that the stimulus had a visual angle of 35×26 deg. The presentation was organised using Psychtoolbox [10], and a set of sequences consisting of 56 photos, were shown for a duration of 6 minutes in total. Three trials were conducted in which different sets were shown to each subject, followed by short breaks.

The subjects, who had sufficient visual acuity, were 6 male university students aged between 19 to 23 years old. The contents of the experiment were explained to all participants in advance, and informed consent was then obtained.

3.2. Measurement Procedure

Both 3ch EEGs and electro-oculograms (EOGs) were recorded synchronously, in response to visual stimuli. Three scalp electrodes for EEGs were positioned in the Frontal (Fz), Central (Cz) and Occipital (Oz) areas, according to the International 10-20 system. The responses generated were measured using a bio-amplifier (ADInstruments: PowerLab4/30, ML13) and recorded as signals on a PC, at a sampling rate of 400Hz, using a low pass filter of 30Hz and a high pass filter with a time constant of 0.3sec. EEG signals were filtered using a band pass filter of 2.0 to 30Hz.

A pair of electrodes for EOGs were placed on both temples and a second pair were placed above and below the left eye. The EOG signals were processed to extract eye movement data as shown in Figure 2. Temporal differentials were calculated in order to reduce signal drift and artefacts such as blinks [11], then eye movements were summarised according to their differences. Saccadic eye movements have been extracted using saccade duration, which is over 20ms. However, the frequency power of eye movements was calculated using raw data. When the EOG level was over a threshold, 80 µV, the trials were excluded from the following analyses since the observation included blinks.
3.3. Subjective evaluation of facial images

To avoid individual differences and misperception during recognition of facial expressions, all subjects were asked to evaluate the facial expression of every photo. The impressions that the facial images produced were measured using a scale called the “Affect Grid” [12], which is used to measure facial expressions [3, 13] and which consists of a two dimensional (9 x 9 point) scale displaying “Pleasant - Unpleasant Feelings” and “High Arousal - Sleepiness” ratings. All 56 photos were rated after being used in the three viewing sessions.

The results of the rating of all subjects across all facial expressions (336 responses: 56 photos x 6 subjects) are summarised in Figure 3. To extract the rating patterns of viewers, cluster analysis using the WPGMA method was performed and two clusters labeled “Pleasant” and “Unpleasant” formed along the axis of “Pleasant - Unpleasant Feelings” [7]. The facial photos subjects viewed were individually classified into two clusters, and all responses of EEG/ERPs and eye movements for each cluster were summarised in the analysis which follows. The percentage of all rated photos which are included in the “Pleasant” cluster is 42%.

4. Results

4.1. EEG responses for the two clusters

Event Related Potentials. The EEG waveforms between the two clusters were compared chronologically, using event related potentials (ERPs). An ERP is an accumulated waveform which emphasises responses to events, such as information processing [8, 14]. ERPs were measured across all trials for each cluster according to the viewer’s evaluation of the differences in chronological responses, as was shown in previous studies [8, 14].

The EEG waveforms between -100 and 600 ms before stimulus onset were extracted, and the average voltage during the 100ms before stimulus onset was set as a baseline. The ERP waveforms between the two clusters on each electrode are compared in Figure 4. The blue lines represent the ERPs of the “Pleasant” cluster, and the red lines represent the ERPs of the “Unpleasant” cluster. To determine whether there are any differences in facial perceptions using ERPs, a pair-wise, statistical test for ERPs across the two clusters was employed, as was in previous studies [15, 16]. This technique identifies the time spans where 15 consecutive t test values were below the p < 0.05 level (df = 34, 2 clusters x 3 sets x 6 subjects - 2). In regards to this test,
two significant time spans were extracted, and are illustrated in Figure 4 as follows, Fz: 142.5~192.5ms, and Cz: 132.5~195.0ms. On both electrodes, the potentials for “Pleasant” images are significantly higher than those for “Unpleasant” images. These results suggest that facial impressions affect ERPs between 132.5ms and 195.0ms after stimulus onset. At the Oz electrode, the significant time slot is too short and occurs too late, so it should be disregarded. In previous studies, the starting times nearly coincided, as emotional facial expressions affected cortical processing with a similar latency of about 120ms [8].

**Frequency analysis of EEGs.** In addition to the time course evaluations using ERPs, frequency domain analysis of the EEG data was also introduced. Traditionally, frequency components of EEGs are recognised as features of waveforms such as α, β and θ waves. Frequency analysis was applied to all of the trials. The conditions of the presentation sequence and the restrictions of the conventional fast Fourier transform (FFT) require that the data length be based on a power of 2. Thus, frequency components were extracted from every 160ms zone (64 sampling points) between 100ms before and 540ms after stimulus onset. In addition, the Hanning window technique was applied to the FFT analysis. In Figure 4, the mean frequency powers in the first two time zones (-100~60ms, and 60~220ms) are summarised across the two clusters.

The frequency components were calculated in 6.25Hz steps, such as at 6.25Hz, 12.5Hz, 18.75Hz . . . , and these may correspond with conventional EEG frequency ranges, such as the θ, α and β band areas [14]. Though the frequency powers of the two clusters shown in Figure 5 are comparable, the powers at 6.25Hz (θ wave) for the second zone (60~220ms) are higher than for the first zone (-100~60ms), which is the baseline. Stimuli onset may affect the powers in the frequency ranges, while the facial images do not influence the other frequency powers of EEGs on Fz and Cz.

**4.2. Comparison of Eye movements**

**Saccade analysis.** Since saccadic eye movements may reflect the emotions caused by the viewing of facial expressions, the features of eye movements were summarised. Since ERP differences appear after 100ms, only their behaviour from 100 to 600ms after stimulus onset is shown in Figure 4. As a previous study suggested that scan-paths of faces were influenced by the viewing of emotional expressions [5], the features of saccadic eye movements were summarised. The saccades for facial images were observed from 100 to 600ms [6]. In Figure 6, the frequencies of saccades and saccade amplitudes between the two clusters are compared across 12 directions [17]. Though there are significant differences in these indices in a few directions, the contribution of emotional factors was not determined. Also, in a time course analysis there were no significant differences between the two clusters.

**Frequency analysis of eye movements.** To compare the features of eye movements, frequency analysis was employed, as were EEGs. Eye movements are
Table 2. Performance of logistic regressions

<table>
<thead>
<tr>
<th>Features</th>
<th>Dim</th>
<th>AIC</th>
<th>$R^2$</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERP(Fz)</td>
<td>18</td>
<td>51.9</td>
<td>0.75</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>ERP(Cz)</td>
<td>25</td>
<td>51.9</td>
<td>0.75</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>ERP(Fz+Cz)</td>
<td>43</td>
<td>51.9</td>
<td>0.75</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>d-FFTpower(Fz)</td>
<td>4</td>
<td>51.9</td>
<td>0.08</td>
<td>70.4</td>
<td>0.70</td>
</tr>
<tr>
<td>d-FFTpower(Cz)</td>
<td>4</td>
<td>51.9</td>
<td>0.30</td>
<td>82.4</td>
<td>0.82</td>
</tr>
<tr>
<td>d-FFTpower(Fz+Cz)</td>
<td>8</td>
<td>51.9</td>
<td>0.41</td>
<td>88.3</td>
<td>0.88</td>
</tr>
<tr>
<td>ERP(Fz) + d-FFTpower(Fz)</td>
<td>22</td>
<td>51.9</td>
<td>0.75</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>ERP(Cz) + d-FFTpower(Cz)</td>
<td>29</td>
<td>51.9</td>
<td>0.75</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>Factor 1 ~ 3</td>
<td>12</td>
<td>49.13</td>
<td>0.57</td>
<td>95.8</td>
<td>0.96</td>
</tr>
<tr>
<td>Factor 1 ~ 3 + d-FFTpower(Fz+Cz)</td>
<td>20</td>
<td>49.13</td>
<td>0.75</td>
<td>100</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Figure 7. (Temporal changes of cross spectra of eye movements (*: $p < 0.05$)).

observed as two dimensional data. The mobility of eye movements can be evaluated as a measure of mental activity using cross spectrum analysis of observed data [18, 19]. Therefore, cross spectra were calculated for each viewing session, as well as for every 160ms time zone of the frequency spectra of the EEGs. Cross spectra were calculated using the MATLAB cpsd function, and the frequency powers are presented in 6.25Hz steps. Mean spectra between the two clusters across all time zones are summarised in Figure 7. The horizontal axis represents frequency, and the vertical axis represents frequency power, in dB. There are significant differences between the two clusters at 18.75Hz (60~220ms), 31.25Hz (220~380ms), and both 37.5Hz and 43.75Hz (380~540ms). The frequency range increases as the duration after stimulus onset increases, though there are no significant differences in the first time zone compared to the baseline.

To summarise these frequency component powers in Figure 7, factor analysis was introduced. Since these powers are correlated with each other, Promax rotation was applied to extract the major frequency components as indices of the reactions. In regards to the factor loading matrix in Table 1, three factor components of frequencies are displayed in bold characters. These mean factor powers are summarised across four time zones in Figure 7, and again the frequency factors of significant differences between the two clusters increases with the duration after stimulus onset. All differences suggest that the degree of two dimensional eye oscillation for “Unpleasant”

Figure 8. Temporal changes of CPSDs of eye movements.
facial images is significantly higher than for “Pleasant” images.

4.3. Prediction for viewed types of facial emotions

Various types of features of biological information, such as ERPs and eye movements, were extracted in order to detect key responses in clusters of “Pleasant” and “Unpleasant” facial emotions. Though most features represent certain differences well, their contribution should be evaluated quantitatively. For this purpose, discriminant analysis is introduced to predict emotional clusters of viewed facial images using the biological responses of viewers. Since both prediction performance and the contribution ratio of the extracted features can be calculated during the analysis, these provide assessments across the types of features. The features are as follows:

- ERP: amplitudes of ERPs from EEGs at the Fz and Cz electrodes during a specific period of time, as shown in Figure 4. These are labeled as ERP(Fz) or ERP(Cz).
- FFTpower: Four components of the FFT analysis for Fz and Cz (6.25, 12.5, 18.75 and 25Hz) may represent the frequency components of their respective EEGs and ERPs, as shown in Figure 5. The differential powers between the first and second zones were used for the analysis. They are labeled as d-FFTpowers(Fz) or d-FFTpowers(Cz).
- Factor scores of CPSD: Three factor scores in the respective EEGs and ERPs, as shown in Figure 3. These are labeled as ERP(Fz) or ERP(Cz).

All indices have been confirmed to represent certain responses between two clusters of facial emotions. Discriminant analysis was conducted for every trial which had 6 subjects.

Analysis of logistic regression as a simple binary classifier [20] was employed to discriminate between the two clusters using the features of the metrics mentioned above. The possibility of a cluster (p) occurring was calculated using the following equation with logit function.

\[ \hat{y}_i = a + b_1 \text{feature}_{1,i} + \cdots + b_n \text{feature}_{n,i} \]

\[ p_i = \text{logit}^{-1}(\hat{y}_i) = \frac{1}{1 + \exp(-\hat{y}_i)} \]

Suffix i represents a trial set of a subject.

Some sets of features were applied to the model. Every model was evaluated for fitness of model using AIC (Akaike Information Criteria), and prediction accuracy was evaluated using \( R^2 \), the AUC (the area under the ROC curve), which is the surface area of the ROC (Receiver Operating Characteristics) curves, as is shown in Table 2. The results show that discriminant performance is perfect when ERP features are employed. The performance with d-FFTpowers features is less than the performance using ERP features. The performance using factor scores of eye movements is less than the scores using ERP features, but in another evaluation using the AIC index the factor scores have the highest performance. Also, combined feature sets using factor scores and d-FFTpowers provide perfect accuracy.

This performance confirms the contribution of raw ERP data, but frequency powers which are extracted from EEG signals at an electrode are insufficient to estimate clusters of viewed facial emotions in comparison with features of ERPs. Factor scores which were extracted from other frequency analyses of eye movement oscillations also represent clusters of viewed facial emotions, and discriminant performance in these cases was better than with powers of d-FFT. This suggests that eye movements reflect the responses of viewed facial emotions sufficiently. The combinations of d-FFTpowers and factor scores of eye movement oscillations improved the performance of discriminant analysis of the clusters of emotions.

In regards to the feature extraction procedures, the features of both d-FFTpowers and eye movements can be represented in each trial, though the features of ERPs need as many trials as possible. As mentioned above, the prediction performances of sets of ERP features and combinations of d-FFTpowers are comparable to the performance of sets of features of eye movement. Therefore, a single trial prediction using these features may be possible. The examination of this will be a subject of our further study.

This result confirms that information from both EEGs and eye movements can be keys to the recognition of clusters of viewed facial emotions.

4.4. Relationships between indices of EEGs and eye movements

In regards to the above mentioned sections, both frequency powers of EEGs and eye movements are caused by the visual perception of facial images. Since the latencies of EEGs and eye movements are different by means of the nature of the reaction, a chronological analysis is required. Chronological correspondence between the two metrics, and recognition of emotional clusters as EEG activity may affect the frequency powers of eye movements. To examine these relationships, correlation analysis was conducted on the frequency powers of EEGs in Figure 5, and on the eye movements in Figure 7.

Three frequency components of EEGs and the third factor of eye movements (Factor3: 31.25~37.50Hz) were analysed, as Factor3 seems to reflect the
cluster differences mentioned in Section 4.2. Frequency components such as \( \theta \), \( \alpha \) and \( \beta \) band areas for Fz, Cz and Oz have been extracted, as stated in Section 4.1. Temporal changes in correlation coefficients between the components of the two clusters mentioned above are summarised in Figure 9 for Fz, in Figure 10 for Cz, and in Figure 11 for Oz. The horizontal axis represents the four time zones, and the vertical axis represents the correlation coefficient. All coefficients are not significant since the number of samples is limited (N=18), however.

In comparing the three figures, the deviations of correlation coefficients in Figure 10 for Cz and in Figure 11 for Oz are relatively large. As Figure 10 for Cz shows, coefficients for both 6.25Hz (\( \theta \) wave area) and 18.75Hz (\( \beta \) wave area) increase monotonically after stimulus onset and throughout the time course. The coefficients for the “Unpleasant” cluster are the highest between 220 and 540ms, while the coefficients for the “Pleasant” cluster decreased between 220 and 540ms. Coefficients of EEGs at 12.5Hz (\( \alpha \) wave area) remained at low coefficient levels.

Figure 11 presents different relationships for Oz, as both coefficients for the “Pleasant” cluster at 12.5Hz (\( \alpha \) wave area) and at 18.75Hz (\( \beta \) wave area) are higher around the onset of stimulus. All coefficients decreased gradually as time zones progressed. Though the coefficients decreased gradually, all coefficients are relatively small, as Figure 9 shows for Fz.

Since the Cz electrode is placed there specifically to detect the activity of the motor cortex, eye movements may reflect the activation. The range of coefficients across the three electrodes was the largest for Cz. In addition, some larger coefficients were observed at Oz in the earliest time zone (-100~60ms). The relationships for the “Pleasant” cluster appeared at around the onset of stimulus. All coefficients were negative for the Fz electrode, and their absolute values increased gradually as the time zones progressed.

Therefore, there are some causal chronologically occurring causal relationships between the activity in the central area, such as at Cz, and eye movements. The results provide evidence that the activity of the motor cortex affects eye movement between 220 and 540ms when “Unpleasant” images of facial expressions are displayed. This phenomenon is more greatly emphasised during the viewing of “Unpleasant” facial expression images than it is during the viewing of “Pleasant” images.

5. Conclusion

This paper confirmed that when images of facial expressions were presented, some metrics of EEGs and eye movements appeared in response to categories of
images of emotional impressions which were “Pleasant” or “Unpleasant”. Various features of EEGs and eye movements were extracted from the participants’ reactions in response to the facial images viewed. The features of typical responses to the two emotional categories were compared, and the chronological relationships between EEGs and eye movements measured using EOGs were examined using time domain and frequency domain analyses in response to “Pleasant” and “Unpleasant” facial images.

In the results, the potentials of ERPs at Fz and Cz for “Pleasant” images were significantly higher than the other potentials, while the powers of frequency components of EEGs for “Unpleasant” images were significantly higher than for the others. The possibility of predicting the emotional category of images of facial expressions was examined using sets of feature data of EEGs and eye movements, and a logistic regression technique. The contribution of these features to the estimation of the proper emotional category was also evaluated. Regarding chronological analysis, the correlation coefficients in frequency powers between EEGs at Cz and eye movements were monotonically increased throughout the time course when “Unpleasant” images were presented. This phenomenon may be explained by eye movements which are stimulated by the motor cortex when it is activated by the viewing of “Unpleasant” images.

These results show that the viewer’s emotional state of mind can be presented as responses to brain activity and viewing behaviour. Therefore, it may be possible to observe a viewer’s emotional state by monitoring EEGs and eye movements. Also, various types of observations are possible using the synchronised relationships of EEGs and eye movements. These observations are fundamental to observing human emotional behaviour.

Acknowledgement
This research was partially supported by the Japan Society for the Promotion of Science (JSPS), KAKEN (B-26282046).

This paper is an extended version which is based on report at BIOSIGNAL2015 (Yasuda et al., 2015[7]) and ETRA2016 (Nakayama and Yasuda, 2016[9]). The authors would like to thank those who provided useful comments regarding these papers.

References
