

Prediction of Two-dimensional Impressions of Images of Facial Emotions using features of EEGs

Satoru Waseda¹ and Minoru Nakayama^{1,*}

¹Tokyo Institute of Technology, Japan

Abstract

The viewing of categories of facial emotions is predicted using features of viewer's scalp potentials, such as event-related potentials (ERPs) measured during the viewing of pictures of facial emotions. All visual stimuli were rated using two-dimensional emotional scales, and the responses for each viewer were converted into sensitivities using item response theory (IRT). This sensitivity to facial emotions can be predicted using discrimination analysis and the extracted features of ERPs recorded during the viewing of the images. The categories of facial emotions viewed were estimated to a certain level of significance using regression analysis, and the sensitivities predicted for the emotional scales were calculated accurately, as performance depended on the reactions to the images of the emotions. The results showed that categories of facial emotions viewed can be predicted to a level of significance using features of scalp potentials.

Received on 26 December 2019; accepted on 21 January 2020; published on 23 January 2020

Keywords: Facial Expressions, ERPs, Emotion, Prediction, Item Response Theory

Copyright © 2019 Satoru Waseda *et al.*, 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.20-8-2019.162800

1. Introduction

The recognition of facial emotion as a form of human visual information processing has been widely discussed [1]. In particular, individual performance in the perception of facial emotions is often measured and compared [2–4]. The emotional perception process is measured using electroencephalograms (EEG) and other bio-signals [5–7]. Some behavioural analysis suggests that emotional perception activity may promote the reproduction of the certain impressions, such as emotional synchronisation [8]. On the other hand, the emotional impressions viewers perceived have to be determined using their responses, such as by rating using a scale. For the presentation of facial emotions, a two-dimensional survey scale, which is known as an “Affect Grid” and consists of scales for two dimensions: “Pleasant” and “Arousal”, is often used [9]. The performance and validity of the scale have been confirmed in previous studies [10, 11].

However, the rating activities produce some individual differences because each individual's emotional sensitivity may influence the ratings. This issue sometimes

influences the perceptual accuracy of facial emotional expressions. As most viewers respond to the “Pleasant” scale [12], the pictures of facial emotions viewed were accurately classified using features of EEGs measured while the images were viewed [13]. If the rating accuracy for pictures of facial emotions were to be improved, it may be possible to predict the perceived categories of emotions using the features of EEGs. This means that the development of an emotion monitoring procedure using viewer's EEGs could be used to estimate emotional categories under various conditions.

In order to create relationships between the viewer's EEG activity and their perceived impressions of facial emotions, both the EEGs and the rating procedure were analysed carefully, and prediction models were developed to present the associations between the two. In particular, individual differences in perception of facial emotions and rating behaviour should be considered in order to determine the appropriate relationships. This paper focuses on these points using a compensation procedure, such as item response theory. The details of the approach used in this paper will be summarised in a comparison with previous studies in the section which follows. In short, when the scores of emotional scales were estimated using EEGs,

* Corresponding author. Email:Nakayama@ict.e.titech.ac.jp

the emotional categories can be predicted using the estimated scores. In this paper, the following topics will be addressed using experimental data from a previous study [13]:

1. The observer's rating behaviour concerning the characteristics of individuals was analysed, and item response theory was used. Surveyed responses were analysed, and their effectiveness was evaluated.
2. The possibility of estimating the sensitivity of viewer's ratings is examined using potentials attached to the scalp while images of facial emotions are viewed.
3. The possibility of predicting both the characteristics and emotional category of visual stimuli while images of facial emotions are viewed is examined.

The following sections will first briefly summarise the topical papers related to the subject, and then present an experiment which observes the viewer's subjective evaluations of images of facial expressions and records EEGs. In Section 4 are the detailed responses and the viewer's behaviour during the rating activities, with the prediction procedures and their performance following in Section 5. Section 6 is a discussion, and the extracted results are summarised in Section 7.

2. Related works

Event-related potentials (ERPs), which sum up EEG waveforms in response to stimuli, are widely used to analyse chronological information processing of pictures of facial emotions. The fundamental approach is to compare ERP waveforms of two facial emotions, such as emotional and unemotional expressions [14, 15]. Also, activity areas of the scalp on the brain are measured for later analysis by presenting various facial emotions, which activate specific regions of the brain [5, 16]. By considering individual responses, the ratings for the two levels of emotions could be predicted using features of EEGs or ERPs [13]. Since the two levels are relatively different, the possibility of predicting the emotional category should be confirmed.

The perception of facial emotions is often measured using a two-dimensional scale [17] called an "Affect Grid" [9], which is widely used to assess facial emotions [18]. However, the rating depends on psychophysical measuring issues [19], in addition to the attitude of the viewer and other factors such as some ability to recognise emotions [20]. Even with the data set of published facial emotions, perceptual performance depends greatly on the characteristics of the viewers [3]. In order to explain the individual differences, the concept of emotional sensitivity has been introduced. Individual differences in sensitivity are analysed for

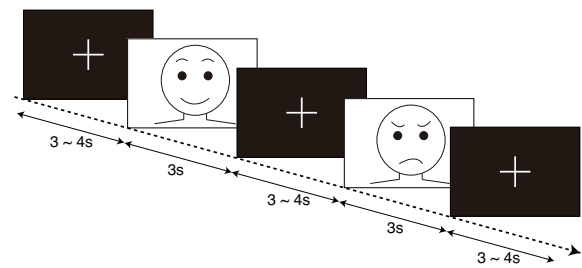


Figure 1. Diagram of stimuli shown.

their effect on the perception of facial emotions [21–23], as this sensitivity may influence the accuracy of the ratings [24].

Since the deviations in rating behaviour are related to a psychometric issues, item response theory (IRT) [25] can be applied. This technique is very useful to compensate for individual bias in ratings [26]. For facial emotion assessment experiments, the IRT model has been introduced in order to extract individual characteristics [27]. This technique can standardise the rating values of the viewers, and their responses can be more easily compared. Also, the characteristics of each individual's rating behaviour are analysed. Therefore, this technique is introduced to the analysis which follows using the data set of the previous experiment [13].

This paper will create relationship models between the viewer's scalp potentials and their emotional categories during the viewing of pictures of facial emotions. If a possible model can be established, viewer's perceived emotional categories could be predicted using the model. The possibility of this will be confirmed in the following sections.

3. Experimental Method

3.1. Stimulus

Images of facial expressions employed as visual stimuli were prepared using the Japanese and Caucasian Facial Expressions of Emotion (JACFEE) collection [28]. 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. Following a black image used to produce eye fixation, each stimulus was displayed for 3 seconds. The subjects were asked not to respond until they had viewed each facial image. 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 [29], and a set of sequences consisting of 56 photos were shown for

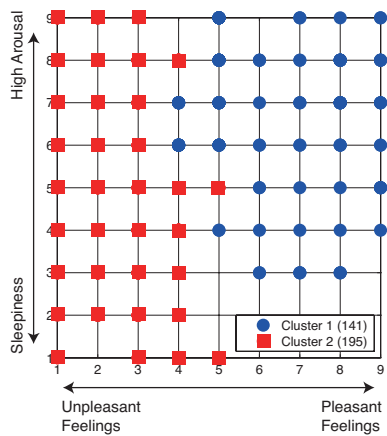


Figure 2. Results of cluster analysis of viewer's responses using the Affect Grid. (Blue circles indicate "Pleasant" cluster and red blocks indicate "Unpleasant" cluster) [30].

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. The contents of the experiment were explained to all participants in advance, and informed consent was then obtained.

3.2. Measurement Procedure

The responses evoked by visual stimuli were measured as scalp potentials, using three electrodes [13]. Three scalp electrodes for measuring EEGs were positioned in the Frontal (Fz), Central (Cz) and Occipital (Oz) areas of the cortex, 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 signal filter with a time constant of 0.3 sec. EEG signals were filtered using a band pass filter of 2.0 to 30Hz. Though subject repeated measure design was employed to determine the factors of the responses, individual differences were also considered. All valid participants were tested across three trials in order to obtain reliable measurements.

3.3. Subjective evaluation of facial images

To avoid individual differences and misperceptions during the 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 an "Affect Grid", which consists of a two dimensional (9 × 9 point) scale displaying "Pleasant - Unpleasant Feelings" and "High Arousal -

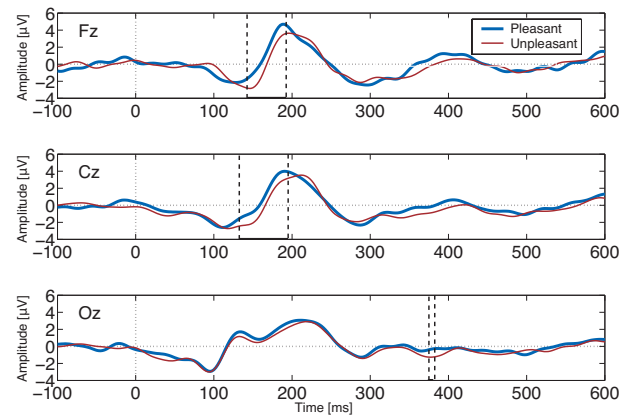


Figure 3. ERP comparisons between two clusters of Electrodes Fz, Cz and Oz, with a dotted line box indicating the time spans where there are significant differences between the two ERPs ($p < 0.05$) [30].

Sleepiness" ratings. All 56 photos were rated after being used in the three viewing sessions.

The results of the ratings of all subjects across all facial expressions (336 responses: 56 photos × 6 subjects) are summarised in Figure 2 [13]. To extract the rating patterns of the viewers, cluster analysis using the WPGMA method was performed, and two clusters labeled "Pleasant" and "Unpleasant" were formed along the axis of "Pleasant - Unpleasant Feelings" [30]. The percentage of all rated photos rated which are included in the "Pleasant" cluster is 42%.

The ERPs generated between the two clusters on the three electrodes are compared in Figure 3. There are significant differences in potentials between the two clusters on electrodes Fz and Cz at the early stage of perception, around 150ms ~ 200ms [13]. This result shows evidence that the ERPs respond to images of facial emotions.

3.4. Item response theory

An IRT model was applied to the ratings of two-dimensional emotional impressions using an "Affect Grid" scale.

$$P_i = \frac{1}{1 + \exp(-Slope * (x + th_i))}$$

Here, P_i is the probability to obtaining a score, which is higher than grade i for sensitivity x , which consists of two parameters: $Slope$ and th_i . All IRT parameters in the above equation were estimated using the IRT package as a graded rating model (GRM) [31]. The parameters, $Slope$ and th_i , suggest the indices of distinctiveness and the difficulty of assessment,

Table 1. Estimated IRT parameters for “Pleasant” scale

Subjects	Slope	th ₁	th ₂	th ₃	th ₄	th ₅	th ₆	th ₇	th ₈
sub1	4.76	-1.27	-0.57	-0.14 [#]	0.33	0.95	1.09	1.48	-
sub2	3.98	-1.34	-1.06	-0.28	0.40	-	-	1.28	1.50
sub3	3.17	-1.45	-0.77	-0.23	0.24 [#]	1.04	1.18	1.54	2.02
sub4	5.11	-	-0.96	-0.45	0.22 [#]	0.92	1.28	1.50	2.07
sub5	3.18	-2.00	-0.94	-0.42	0.15 [#]	0.67	1.13	1.72	2.70
sub6	3.57	-0.93	-0.63	-0.10 [#]	0.13 [#]	0.46	0.60	1.04	2.15

[#]: not significant ($p > 0.10$)

Table 2. Estimated IRT parameters for “Arousal” scale

Subjects	Slope	th ₁	th ₂	th ₃	th ₄	th ₅	th ₆	th ₇	th ₈
sub1	3.77	-	-	-2.17	-1.19	-0.51	-0.19 [#]	0.29	1.09
sub2	1.18	-3.33	-2.68	-1.67	-0.53	0.07 [#]	0.95	1.72	2.88
sub3	2.65	-2.67	-2.28	-1.25	-0.73	-0.32	0.32	0.73	1.27
sub4	2.78	-	-2.69	-1.47	-1.35	-0.78	-0.18 [#]	0.74	1.70
sub5	2.34	-	-1.97	-1.29	-0.65	-0.06 [#]	0.94	1.57	-
sub6	1.23	-2.49	-1.79	-	-0.39	0.07 [#]	0.60	1.23	2.02

[#]: not significant ($p > 0.10$).

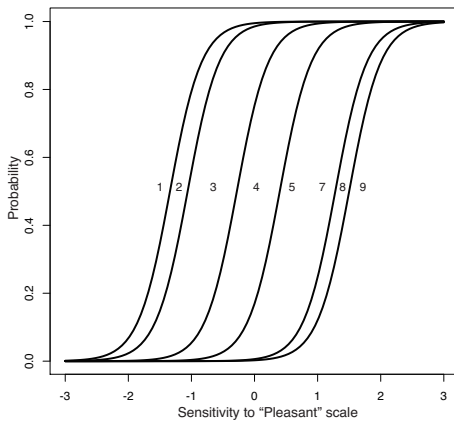


Figure 4. An example of operating characteristic curves (OCCs) for sub2 in Table 1.

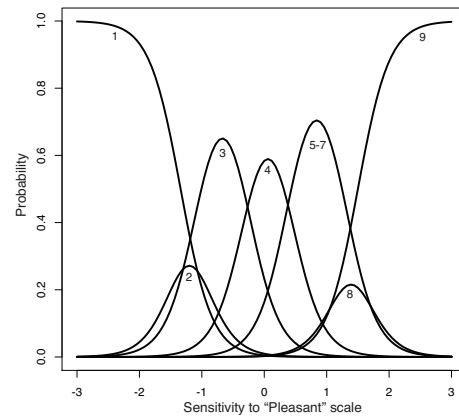


Figure 5. An example of category response curves (CRCs) using the probabilities for sub2 of “Pleasant” ratings in Figure 4.

respectively. They are estimated using an expectation-maximum (EM) algorithm. In this paper, the SAS procedure was applied [31].

4. Results

4.1. Individual parameters for IRT models

The details of the viewer’s rating behaviour for facial emotional expressions are evaluated using the IRT model mentioned above. The parameters of the IRT equation for each participant were estimated using a piece of software [31]. These parameters are summarised in Tables 1 and 2. The missing grades are not given parameters. In the results of significance

tests for all parameters, some tests were not significant. These parameters show the rating behaviour, and suggest that rating sensitivity exists.

This is summarised as operating characteristic curves (OCCs), as shown in Figure 4. If the spaces between the two curves are narrow, the two ratings may not be easily discriminable.

In order to illustrate the difficulty of rating emotional images using the scale, the category response curves (CRCs) which display the probabilities of rating sensitivities are calculated as the differences between the two curves in Figure 4. The results are calculated as $1 - P_1$ and $P_{i-1} - P_i (i = 2, \dots, 9)$ and summarised in Figure 5. As all peaks for curves show rating values, the

rates for “2” and “8” have low probabilities, and they are varied by the neighbouring curves. Therefore, these ratings may not be effective.

4.2. Prediction of ratings for facial emotions using EEGs

As mentioned in the introduction, the possibility of predicting two-dimensional rating values using the “Affect Grid” scale should be examined using the following features of EEGs and ERPs. Four frequency powers, 6.25, 12.5, 18.75 and 25.0Hz, were calculated for four 160ms durations from 100ms before to 540ms after stimulus onset using electrodes Fz, Cz and Oz. The prediction technique employed used a logistic regression procedure noted as follows:

$$\hat{y}_i = a + b_1 \text{feature}_{1,i} + \dots + b_n \text{feature}_{n,i}$$

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

The $\text{feature}_{n,i}$ consists of the frequency power, as mentioned above. Suffix i represents one subject’s trial. The performance of each subject was evaluated using a subject leave-one-out procedure which measured the performance of a subject using a model consisting of the remaining subjects.

Predictions based on Single-trial features. The 9-point scale ratings were predicted using frequency powers for 3-channel EEGs. In order to consider individual differences, prediction functions were produced for each subject. The analysed data consisted of 168 sets (single trials: 56 photos \times 3 sessions) with 48 features (frequency band: 4 \times electrodes: 3 \times time zone: 4) for each subject. The prediction models were evaluated using a subject leave-one-out procedure.

The prediction performance remained around the level of chance (accuracy = 12.89%), however. During the analysis, two major points were extracted as causes of the reduced performance, such as ambiguous rating behaviour and deviations in features. As mentioned in the previous section, the sensitivity of the rating scale is sometimes relatively low. In regards to the overall features of the category response curves for all subjects, the rating scales should be reduced from 9 to 4 points. Also, single-trial EEGs contained a lot of noise, so that ERPs were calculated for the identical pictures over three sessions.

Predictions based on ERPs for 3 trials. The picture ratings were converted into a 4 point scale to better reflect individual responses on a scale. The conversion considered the sizes of the areas in the category response curves in Figure 5. As a result, Figure 6 shows an example for a case of sub2, with three new

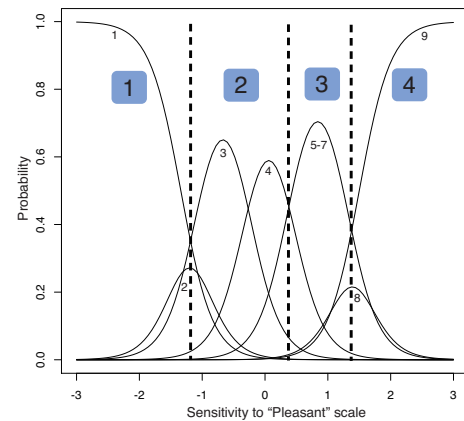


Figure 6. An example of a modified scale which considers CRCs in Figure 5.

Table 3. Prediction performance for “Pleasant” scale

Converted Rating	Prediction				Total
	1	2	3	4	
1	13	27	3	0	43
2	44	110	27	2	183
3	1	20	36	4	61
4	0	0	6	43	49
Total	58	157	72	49	336

criteria and four levels applied to rate sensitivities. The ERPs were calculated from 3 sessions for each image of a facial emotion, and prediction analysis of the four levels was then applied to 56 sets (ERPs for 3 trials) with 48 features. The logistic regression technique was also applied, and the prediction was based on the highest probability of the four levels. All features were employed in the prediction.

Prediction performance was evaluated using a subject leave-one-out procedure which predicted the responses of each subject using a model trained with response data of the remaining subjects.

The results for the four levels in “Pleasant” is summarised in Table 3. The rows indicate 4-point scale ratings and the columns indicate predicted ratings. Therefore, the diagonal cells show correct predictions. The rating is an ordered scale, and the rank correlation coefficients can be calculated from these. Though the frequencies for lower or higher ratings, such as “1” or “4” are lower, certain numbers of predictions of these were observed. The rank correlation coefficients are $r = 0.72$ for the “Pleasant” scale and $r = 0.61$ for the “Arousal” scale. Individual prediction performance is summarised in Figure 7. The green plot indicates mean performance, which was 60% for “Pleasant” and 51% for “Arousal”, and the error bars indicate the standard deviation.

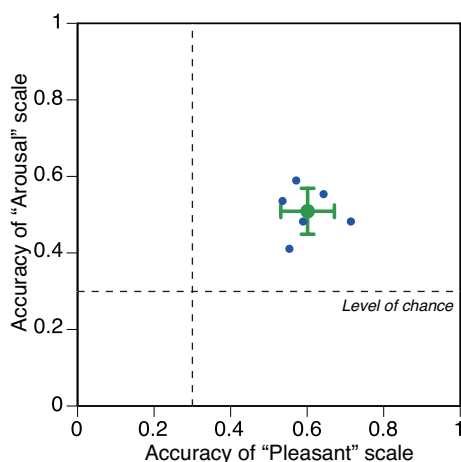


Figure 7. Prediction performance of emotional scales for “Pleasant” and “Arousal”, and their means.

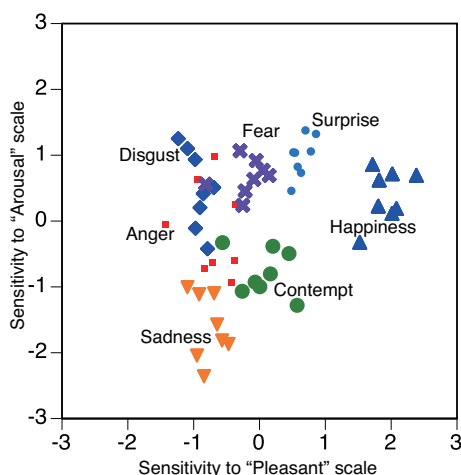


Figure 8. Mean IRT factor scores from “Affect Grid” rating scores in Figure 2

The result shows that both emotional scales can be estimated using features of ERPs, when the responses to the scales are characterised using IRT models.

5. Prediction of characteristic values for each facial image

5.1. Sensitivity to the pictures presented

As mentioned in the section above, the IRT procedure can provide rating scores which consider the characteristics of individuals. Mean scores of sensitivity for both the “Pleasant” and “Arousal” scales of each image, which were rated by participants during the experiment as shown in Figure 2, summarise sensitivities using IRT analysis, and are shown in Figure 8. This figure illustrates that the emotional states in the pictures presented

are almost always reproduced using the sensitivities, and the 8 pictures of each specific emotional category are located in positions surrounding each of the emotions. Therefore, participants might recognise most of the facial emotions in the pictures presented, and rate them according to their individual bias.

In comparison with previous studies such as Russel & Bullock [17], positive ratings for the “Arousal” scale were suppressed. Some factors in this experiment might have influenced the impressions of viewers. Though several facial emotions, such as “Anger”, “Disgust”, “Fear” and “Surprise”, are located in neighbouring positions of the scales [17], the discrepancy seems unclear because the amplitudes of both scales are relatively small. Another previous study suggested that the recall rate for emotions is almost as high as it is for emotional perception (“Happiness”:98.3%, “Surprise”:92.3%, “Fear”:54.6%, “Disgust”:74.7%, “Anger”:64.2%, “Sadness”:71.9%, “Contempt”:76.7%) [3]. This performance may reflect the speed and accuracy of perception of facial emotions [2, 4, 32].

In this section, the possibility of predicting the sensitivity of each picture of a facial emotion is discussed using observed features of EEG signals during viewing.

5.2. Prediction of sensitivity values from EEG signals during viewing

Prediction procedure. Since the perception of facial emotions may affect viewer’s EEG responses, the relationships between features of EEG waveforms and the sensitivity towards emotional impressions when viewing pictures of facial emotions are analysed. The procedure for predicting the sensitivities of both “Pleasant” and “Arousal” from features of EEG signals was developed using logistic regression analysis of both of the sensitivities and the sets of EEG features for each photo. Each viewer rated the 56 pictures of facial emotions using the “Affect Grid”. Mean sensitivities of 56 pictures of facial emotions are estimated using features of EEGs of viewers, as shown in Figure 8. The features of EEGs are extracted using the following procedures, as shown in Figure 9:

1. The ERP waveforms of all viewers from every trial were summed up (56 ERP waveforms).
2. Both Fz and Cz were selected as the targeted electrodes since emotional reactions were confirmed in Figure 3.
3. Eight frequency power components from a certain point (160ms) were extracted from the features of ERP waveforms using an overlapping technique, from 2.5 to 20Hz in 2.5Hz steps.

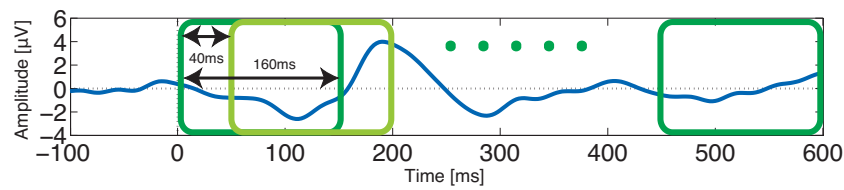


Figure 9. ERP feature extraction procedure: 160ms windows move stepwisely by 40ms.

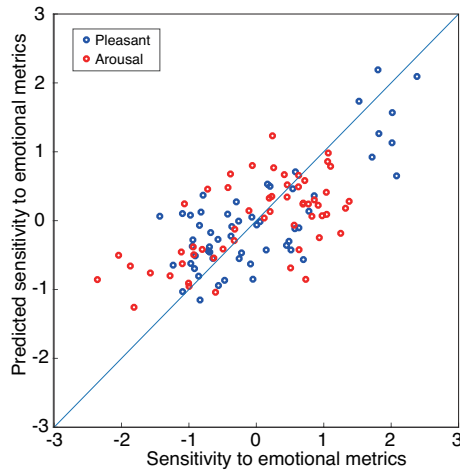


Figure 10. Relationship between Sensitivity to emotional metrics and Predicted values.

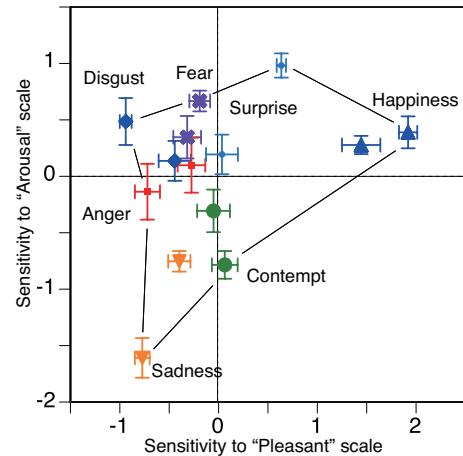


Figure 11. Measured and predicted sensitivities for seven emotional categories (Error bars indicate Standard Errors).

4. Since some delayed responses to images of facial emotions were observed, the observation time window at 160ms moved step-wisely from 0ms-160ms to 440ms-600ms in 12 steps, in order to cover temporal changes in the wide range of images, as Figure 3 shows.
5. As a result, 192 dimensional features ($2\text{ch} \times 8$ frequency powers \times 12 time zones) were extracted from ERPs for 56 pictures of facial emotions.

The sensitivity values were also predicted using a logistic regression function which is mentioned in the section above. In order to extract the key features of the estimation, the optimisation of feature selection was conducted using a stepwise procedure.

Prediction results. As a result of the optimisation process, two regression functions for each sensitivity toward facial emotions were produced. For the values of the “Pleasant” scale, the function consisted of four features ($R^2 = 0.58$): three features came from 2.5, 5, and 12.5Hz on the Cz electrode at 240 ~ 400ms, and one feature came from 7.5Hz on the Cz electrode at 160 ~ 320ms. For the values of the “Arousal” scale, the function consisted of five features ($R^2 = 0.37$): two features came from 2.5 and 10Hz on the Fz electrode at

280ms ~ 440ms and 440ms ~ 600ms respectively, and three features came from 10, 15, and 20Hz on the Cz electrode at 40ms ~ 200ms, 240 ~ 400ms, and 280ms ~ 440ms respectively.

Most features were selected from around 200ms ~ 400ms, which might be an image processing phase after the perception of emotional images was detected in Figure 3. Also, most features came from the Cz electrode. In regards to the 9 ERP features of each picture viewed, two-dimensional sensitivities were predicted using the two functions above.

These predicted values are compared with the original values as an index of sensitivity to emotional metrics, in Figure 10. They are strongly correlated with each other, and the correlation coefficient for “Pleasant” is 0.76, and 0.61 for “Arousal”. Therefore, the prediction functions almost exactly reproduced the sensitivity values of the features of ERPs during the viewing of images of facial emotions. Though the regression line is almost a diagonal one, the range of the predicted values is more narrow than the deviation of the original values (“Pleasant”:[-1.4,2.3] \rightarrow [-1.1,2.1], “Arousal”:[-2.3,1.3] \rightarrow [-1.2,1.2]).

Since these two-dimensional sensitivities can suggest facial emotions, as Figure 8 shows, the facial emotion prediction performance can be examined. The means of

values of sensitivity for the 8 pictures in each category of emotion are summarised along with their standard errors, in Figure 11. These means are connected by lines to create a polygon showing facial emotion sensitivities. As the means of sensitivity for “Happiness” and “Sadness” are relatively large, these facial images can be recognised more easily. Additional means of sensitivity are calculated using the predicted values which are illustrated in Figure 10. All plots for means of predicted values are located inside the polygon. As Figure 11 shows, some emotions are located around the central position of the “Affect Grid”, which suggests a non-emotional, or neutral impression. In particular, the four emotions “Surprise”, “Fear”, “Disgust”, and “Anger” almost over-lap each other. Though the sensitivity values were almost all predicted using the optimised regression functions, and were correlated with the original values, the prediction performance between facial emotions was different, as Figure 11 shows. Because the estimation conducted was based on ERP waveforms, some key pieces of information can be detected at other points on viewer scalps.

This result suggests that the prediction of sensitivities of facial emotions cannot be estimated for all images, though the overall tendencies can almost all be reproduced. The improvement of these insufficient estimation procedures will be a subject of our further study.

6. Discussion

This paper examines the possibility of predicting two-dimensional scales of emotional impressions during the viewing images of facial emotions, using several scalp potentials of the viewers. Since EEG responses during the viewing of facial images may reflect the viewer’s facial image processing process, these activities must suggest that viewers recognised the emotional impressions they saw. In general, recording EEGs requires many electrodes, such as 16ch or more, based on the international 10/20 system, and the measurement of scalp responses during every-day life is not easy. Therefore, this study conducted a feasibility study of prediction of emotional scales of images of facial emotions viewed using only several scalp responses, where these potentials were measured using just two or three electrodes. As the previous studies suggested, individual differences always influenced most of the metrics. In Section 4, the individual differences in rating behaviour were confirmed, and the deviations in responses due to individual characteristics were compensated for using IRT models. As a result, individual accuracy for both emotional scales was significant when the prediction models were used, and the deviations between subjects were suppressed, as Figures 7 and 8 show.

The estimation of the 7 emotional categories of both emotional scales is possible using ERP signals during the prediction phase, but the accuracy is insufficient for some emotions, as Figure 11 shows. Since the ERP signals were based on two electrodes, these limited information resources might influence performance. Because this feasibility study was conducted using a small number of participants, the validity of the results and the ability to use the procedure should be confirmed with a larger number of participants.

The final goal of this study is to predict a viewer’s emotional condition using simple biological information and images of goods and services. In a sense, this preliminary work confirms the possibility of the plan. In order to develop a more detailed examination process, additional biological information and the individual differences of the many participants should be examined. These studies will be the subject of our further study.

7. Conclusion

In order to estimate viewer’s emotional states in response to viewing images of facial emotions, the possibility of estimating viewer’s perceived rates of emotional impressions in two-dimensions were examined using their EEGs recorded while viewing the facial emotion visual stimuli. In the results, the following points were discovered.

1. Viewer’s rating behaviour, which was based on two-dimensional emotions such as “Pleasant” and “Arousal” were analysed using an “Affect Grid” which involved applying item response theory (IRT) in order to compensate for individual characteristics in the perception of images of facial emotions. It was confirmed that the intensities which were calculated represented the categories of facial emotions viewed almost exactly.
2. Using the extracted features of ERPs, which were summed up using EEGs from the three viewings of each individual and picture, viewer’s rating sensitivities can be estimated accurately using discrimination analysis based on the IRT analysis.
3. Using the features of ERPs in a time series of EEGs from all pictures, the characteristics of each facial emotion visual stimulus were predicted correctly. The facial emotions viewed were predicted almost exactly using estimated values for characteristics based on regression analysis. Prediction performance depended on the category of emotion, with the highest performance for “Happiness” and the lowest performance for “Fear”, “Disgust” and “Anger”.

The recognition of facial emotions is based on human image processing and emotion sensitivity, and therefore wider and more accurate collection of information from scalp activity may be required. The improvement of performance and the introduction of more effective feature extraction procedures for more accurately predicting emotional expressions will be a subject of our further study.

Acknowledgement

This research was partially supported by the Japan Society for the Promotion of Science (JSPS), KAKEN (17H00825).

References

- [1] BRUCE, V. (1988) *Recognising faces* (London, UK: Lawrence Erlbaum Associations Ltd.).
- [2] KIROUAC, G. and DORÉ, F.Y. (1984) Judgement of facial expressions of emotion as a function of exposure time. *Perception and Motor skills* **59**: 147–150.
- [3] BIEHL, M., MATSUMOTO, D., EKMAN, P., HEARN, V., HEIDER, K., KUDOH, T. and TON, V. (1997) Matsutmoto and Ekman's Japanese and Cacasian facial expressions of emotion (jacfee): Reliability data and cross-national differences. *Journal of Nonverbal Behavior* **21**(1): 3–21.
- [4] SCHERER, K.R. and SCHERER, U. (2011) Assessing the ability to recognize facial and vocal expressions of emotion: Construction and validation of the emotion recognition index. *Journal of Nonverbal Behavior* **35**: 305–326.
- [5] BATTY, M. and TAYLOR, M.J. (2003) Early processing of the six basic facial emotional expressions. *Cognitive Brain Research* **17**: 613–620.
- [6] 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.
- [7] 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.
- [8] SUZUKI, A. (2014) Facial expression recognition and embodied simulation. *Japanese Psychological Review* **57**(1): 5–23.
- [9] RUSSELL, J.A., WEISS, A. and MENDELSON, G.A. (1989) Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology* **57**(3): 493–502.
- [10] 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.
- [11] 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.
- [12] BARRETT, L.F. (2006) Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review* **10**(1): 20–46.
- [13] NAKAYAMA, M. and YASUDA, M. (2019) Feature characteristics of erps and eye movements in response to facial expressions. *EAI Endorsed Transactions on Context-aware Systems and Applications* **18**(15)(e3): 1–9.
- [14] EIMER, M. and HOLMES, A. (2002) An erp study on the time course of emotional face processing. *NeuroReport* **13**: 1–5.
- [15] HOLMES, A., NIELSEN, M.K. and GREEN, S. (2008) Effects of anxiety on the processing of fearful and happy faces: An event-related potential study. *Biological Psychology* **77**: 159–173.
- [16] EIMER, M. and HOLMES, A. (2007) Event-related brain potential correlates of emotional face processing. *Neuropsychologia* **45**: 15–31.
- [17] RUSSELL, J.A. and BULLOCK, M. (1985) Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology* **48**(5): 1290–1298.
- [18] ADOLPHS, R. (2002) Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews* **1**(1): 21–62.
- [19] MARNEWECK, M., LOFTUS, A. and HAMMOND, G. (2013) Psychophysical measure of sensitivity to facial expression of emotion. *Frontiers in Psychology* **4**(63): 1–6.
- [20] SCHLEGEL, K., PALESE, T., MAST, M.S., RAMMSAYER, T.H., HALL, J.A. and MURPHY, N.A. (2019) A meta-analysis of the relationship between emotion recognition ability and intelligence. *Cognition and Emotion* **34**(1): 1–23.
- [21] GAO, X., CHIESA, J., MAURER, D. and SCHMIDT, L.A. (2014) A new approach to measuring individual differences in sensitivity to facial expressions: influence of temperamental shyness and sociability. *Frontiers in Psychology* **5**(26): 1–9.
- [22] FISCHER, A.H., KRET, M.E. and BROEKENS, J. (2018) Gender differences in emotion perception and self-reported emotional intelligence: A test of the emotion sensitivity hypothesis. *PLoS ONE* **13**(1): 1–19.
- [23] RUTTER, L.A., SCHEUER, L., VAHIA, I.V., FORESTER, B.P., SMOLLER, J.W. and GERMINE, L. (2019) Emotion sensitivity and self-reported symptoms of generalized anxiety disorder across the lifespan: A population-based sample approach. *Brain and Behavior* **9**(6): 1–7.
- [24] LYUSIN, D. and OVSYANNIKOVA, V. (2016) Measuring two aspects of emotion recognition ability: Accuracy vs. sensitivity. *Learning and Individual Differences* **52**: 129–136.
- [25] LINN, R.L. (1989) *Educational measurement, 3rd edition* (New York, USA: Collier Macmillan Publishers).
- [26] UTO, M. and UENO, M. (2018) Empirical comparison of item response theory models with rater's parameters. *Heliyon* **4**: 1–32.
- [27] SUZUKI, A., HOSHINO, T. and SHIGEMASU, K. (2006) Measuring individual differences in sensitivities to basic emotions in faces. *Cognition* **99**: 327–353.
- [28] MATSUMOTO, D. and EKMAN, P. (1988), Japanese and Caucasian facial expressions of emotion (JACFEE) and neutral faces (JACNeuF). San Fransisco State University, San Francisco, CA, USA.
- [29] BRAINARD, D.H. (1997) The Psychophysics Toolbox. *Spatial Vision* **10**: 433–436.

- [30] YASUDA, M., DOU, Z. and NAKAYAMA, M. (2015) Features of event-related potentials used to recognize clusters of facial expressions. In *Proceedings of International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNAL 2015)*: 165–171.
- [31] SAS, SAS/STAT 14.1 User's Guide: The IRT Procedure. URL <https://support.sas.com/documentation/onlinedoc/stat/141/irt.pdf>.
- [32] COLONNELLO, V., RUSSO, P.M. and MATTAROZZI, K. (2019) First impression misleads emotion recognition. *Frontiers in Psychology* **10**(527): 1–6.