A DISCRIMINATION ALGORITHM OF EVENT RELATED POTENTIALS EVOKED BY CHARACTERS

M. F. Funada¹, S. P. Ninomija²
¹Department of Business Management, Hakuoh University, Oyamashi, Tochigi, JANAN
²Faculty of Science and Engineering, Aoyama Gakuin University, Sagamihara, Kanagawa, JAPAN

Abstract
In this paper, the authors define an EEG evoked by visualizing a character as a kind of Event related Potentials, and try to estimate the character from the Event Related Potentials (ERPs). The approaches to the purpose are to design an experiment at first, to measure ERPs using an adaptive filter, and to find the characteristics of ERPs corresponding to the characters used as stimuli. Next, we consider an algorithm to estimate the characters causing the ERPs. Finally we discuss the stability and reliability of the algorithm.

Key Words
caracter, EEG, ERP, discrimination, visualize, vowel

1. Introduction
Brains control human beings, and human beings consider various things in the brains using a lot of languages. Electroencephalogram (EEG) is an identifier to show the states of brains and it is natural to consider a special EEG appears when a human being visualizes a special word in mind. Since the special EEG is tightly linked to the event that a human being visualizes a word, the EEG is considered a kind of Event Related Potentials. So, the authors define the EEGs as a kind of Event Related Potential (ERP), and try to find the special ERPs.
In order to make the problem simpler, we focus on the sub-problem that estimates not words but characters from ERPs. That is, our hypotheses are three; the first one is that different characters evoke different ERPs, the second is that the relation is one to one correspondence, and the last is that the visualized characters can be estimated from the ERPs.
The authors tried to make sure that the first and the second hypothesis about Japanese vowels[2,3]. In this paper, the authors tried again to make sure that the both hypotheses are correct and challenged to verify the correctness of the last hypothesis. The methods are that we executed several experiments, found the characteristics of ERPs showing the visualized characters, and considered a discrimination algorithm of ERPs. If it is possible to estimate the characters with high quality, the application area may be wide, and wearable computers become realizing truly.

2. Methodology
2.1 Experiments
The experiment to obtain ERPs is followings;
1) Subjects; ten persons whose ages are 19-22 years old. We use “a”, “b”, ..., and “j” to identify them.
2) Place for the experiment; one of authors’ laboratory.
3) Stimuli; five white and black circles. We call the repetition of six circles “one term”. We repeat ten terms for a task and call the experiment “a set”. Three sets with different tasks are repeated for all the subjects with two seconds rest between the sets.
4) Display of the circle; they are displayed on a 19inch CRT in front of a subject.
5) Task for the subjects; the subjects visualize the following three kinds of Japanese characters.
I. Japanese vowels in Fig.1. We use a characters’ set {A, I, U, E, and O} for representing the vowels.
II. A set of Japanese characters {A, O, I, U, and MI} which mean “blue sea”.
III. A set of another Japanese characters {A, KA, I, I, and E} which mean “a red house”.
When the subjects watch the first white circle on CRT, they visualize the first character in I, II, or III. By the next white circle they visualize the next character in I, II, or III. After they visualize all characters, the black circle is displayed. They know an end of a term and do nothing(a short break). Only the subject “a” executes

![Fig.1 The visualized characters by the circles as stimuli](image)

“E” “O” “a short break”
the task I five times to know the proper number of average for obtaining an ERP.
6) Period for displaying circles; one second for every white and black circle. An interval between circles is at random with the range [800msec, 1200msec].
7) Length of experiment; about two minutes are required for a set, and two minutes' short rest between the sets is also required. So an experiment takes ten minutes totally.
8) EEGs; single polar eight channels of “10-20 methods” are used for measuring Electroencephalogram (EEG).
   The positions of measurement are Fp1, Fp2, C3, C4, O1, O2, Cz, and Pz. The base is A1 connected with A2.
9) The sampling frequency for A/D; one kHz.

2.2 Analysis

1) Recorded EEGs are divided into six groups corresponding to the characters and the black circle.
2) The EEGs of all groups are filtered by a kind of adaptive filter which is made by one of authors.
3) The filtered data are normalized using their averages and standard deviations.
4) The normalized data in the group are averaged to obtain an ERP evoked by visualizing characters. The data in the other groups are so on.
5) The averaged data are represented as several kinds of graphs to show the characteristics of the ERPs.

3. Results

3.1 The recorded data and filtered data

Fig.2 is an example of recorded data. The subject is “b” and she is visualizing “I”. The horizontal axis is the time after a circle is displayed, and the vertical axis is the amplitude. There is a possibility of mixture noises like alternating current with 50Hz, eye blinking, electromyogram and so on.
Fig.3 is the filtered data of Fig.2. Noises with higher frequencies than EEG component like alternating current are decreased in the data.
Fig.4 is the normalized data of Fig.3. The range of each data is almost the same. The normalization makes the influence of each data to the averaged data the same. This processing has an effect to decrease the influence of the large noises to averaged data.

3.2 The relationship between ERPs and visualized characters

Fig.5, 6, 7, 8, and 9 are one of examples of fifty times averaged ERPs under the task I measured at the position C3. The horizontal line is the time after a circle is displayed and the vertical line is normalized amplitude of ERP. The bold lines are ERP's themselves and thin lines mean the values “(average)-(standard deviation)” and “(average)-(standard deviation)”. The arrows in these figures show the positions with small standard deviations. They are considered as the points including signals because a standard deviation is considered as an identifier to show the “single to noise ratio (S/N)” and the points with small (S/N) include large signal.
Fig.10 is another expression of the fifty averaged ERPs. The positive peak around 150msec and the negative peak around 200msec are almost the same each other. The peaks are considered “visual evoked potentials (VEPs)” caused by watching the stimuli (circles). After 200msec, there are several peaks and they are different each other. The peaks appearing after 200msec show the responses of visualizing characters because the subjects recognize the circle and then visualizing characters.
Fig.11 shows the variances about ten times and fifty times averaged ERPs. The variances change largely through fifty times’ trials, and the repetition of recordings
concentration and stabilities of background activity of brain.

3.3 Ten averaged ERPs

Fig.12 is an example of ten times averaged ERPs obtained from the subject “b”. The marks from “A” to “O” mean the kinds of visualized characters and the black circle is no task. The waveform in the part indicated by the arrows pulled horizontally on the top of waves resembles alpha waves and it is estimated that no signal is included in the part. So the responses caused by visualizing characters appear after 300msec about the subject “b”. Fig.13, 14, 15, 16, 17 and 18 are ERPs for each character. The variances about “U”, “E”, and “O” are smaller than others. “MI” and “KA” are very different from the vowels.
These figures also suggest that the possibility of estimation of visualized characters from obtained ERPs.

4. Discussions

In this section, we would like to discuss our three hypotheses in the “Introduction”. The first hypothesis is that visualizing different characters evokes different ERPs. We consider the hypothesis using the another ERPs which were measured on October 1. The ERPs evoked by visualizing “U”, “E” and “O” are more stable than the others and Fig. 19, 20 and 21 show them. The arrows point the typical potentials for each ERP. They do not resemble each other, however, we calculate the correlations among them in the period [250msec, 600msec]. Fig. 22 is an example of the correlations. The horizontal axis is the correlation between the ERP caused...
Fig.19 Ten times averaged ERP caused by visualizing “U” (subject:b)
They are recorded on October 1, 2003.

Fig.20 Ten times averaged ERP caused by visualizing “E” (subject:b)
They are recorded on October 1, 2003.

Fig.21 Ten times averaged ERP caused by visualizing “O” (subject:b)
They are recorded on October 1, 2003.

Fig.22 Plots of the correlations (x,y). “x” is the correlation between the first recorded averaged “E” and the other averaged ERP. “y” is also the correlation between the second recorded averaged “E” and the other averaged ERPs.

Fig.23 Plots of the correlation (x, y). “x” is the correlation between the first recorded averaged “E” and the other averaged ERP. “y” is also the correlation between the second recorded averaged “E” and the other averaged ERPs.

by visualizing “E” in the task I and the others, and the vertical axis is the correlation between the ERP caused by “E” in the task III and the others. The plotted points have tree characteristics; the first one is that the ERPs by visualizing “E” resemble each other, the second one is that the other ERPs are almost ordered on a line, and the last is that ERPs by “U” and “O” are closer than those by “A” and “I”. Anyway, ERPs caused by visualizing “E” are clearly different from others.

Fig 23 are the plots of correlations between the averaged ERP by “E” recorded on June 13 and those recorded on October 1. The horizontal axis is the former, and the vertical axis is the latter. In the figure, the phenomena that ERPs caused by “E” resemble each other are reproductive.

They are recorded on October 1, 2003.

Following these rules, we represent ERPs by the following vectors.

Rule1: counting the number of positive peaks around [200msec, 700msec]. We represent the number the variable “m”. The period is slightly changed for individuals.

Rule2: making a vector with “m” dimension.

Rule3: if the amplitude of the first peak is large, then the first component of the vector is “1”, and otherwise “0”. The threshold for dividing into “0” or “1” is determined for individuals each.

Following these rules, we represent ERPs by the following vectors.
\[
ERP_i \Rightarrow (v_{i1}, v_{i2}, \ldots, v_{im}) \\
i: \text{identifier of data} \\
v_{i1}, \ldots, v_{im} : 0 \text{ or } 1
\]

Then we define the distance between the vectors:

\[
dst(i, j) = \sqrt{\sum_{k=1}^{m} (v_{ik} - v_{jk})^2}
\]

\(j\) : \(j\)th ERP evoked by the task \(I\), \((j = 1, 2, \ldots, 5)\)

\(i\) : \(i\)th ERP evoked by the tasks \(II\) and \(III\).

The algorithm for discrimination is following:

\[
\text{if } ERP_j \text{ has } \min_j \{ \text{dst} (i, j) \}
\]

then we estimate \(\text{char} (i) = \text{char} (j)\)

We apply the algorithm to all recorded ERPs, then we obtain the results shown in Fig.24. The correctly detecting ratios distribute from 69.2% to 90.0%. The results show the high possibility that the relation between shapes of ERPs and the visualized characters is “one to one” correspondence, and we can estimate the visualized characters from ERPs.

5. Conclusions

From above trials and discussions, we conclude followings;

1) We define the potentials caused by visualizing characters as a kind of ERPs.
2) We make three hypotheses that the ERPs by visualizing different characters are different each other, the relation between ERPs and characters is one to one correspondence, and we can estimate the visualized character from the ERP.
3) In order to verify the hypotheses, we repeated experiments about three cases; the first is to visualize Japanese vowels sequentially, the second is to visualize a sequence of characters meaning “blue sea”, and the last is to visualize that of characters meaning “a red house”.
4) The ERPs caused by visualizing characters have larger variance about normalized amplitude than that of ERPs caused by ordinary visual or auditory stimuli.
5) The task of visualizing characters is more difficult to keep the concentration of subject high level than ordinary tasks that make an action by a stimulus. The results of the experiment show that the limit keeping concentration of subjects high is around ten times repetition.
6) The difference among individuals recognized on ERPs is also larger than that of ordinary type of stimulus-action tasks.
7) We make sure that the ERPs caused by visualizing “E” are different from others and the ERPs are very stable and reproducible using averaged ERPs and the correlation among them in [250msec, 600msec]. We consider the result explains that the first hypothesis is established.
8) The correlations separate the ERPs caused by “E” from the ERPs by the others, however, they do not separate the other ERPs from each other. And the difference among individuals makes it difficult to measure of ERPs caused by visualizing characters. And the problems also make it difficult to detect the visualized characters.
9) In order to overcome the difficulties, we make the rules to translate the characteristics of ERP to a vector with values “1” or “0”, and consider an algorithm to estimate the characters using the vectors.
10) The reliability of the algorithm suggests the possibility of discrimination of visualized characters through ERPs. We consider that the second and the last hypotheses are also correct.

Finally, we consider that it is further problem for us to find typical ERPs corresponding all characters and it’s also further problem to devise an effective filter for obtaining the typical ERPs. It’s desired that we execute the experiment about more subjects.

References