2011 International Symposium on Nonlinear Theory and its Applications
NOLTA2011, Kobe, Japan, September 4-7, 2011

# Variability in EEG with Single Point Sensing as Inter-Individual Difference Measure Using Self-Organizing Map

Shin-ichi Ito†, Masashi Hamaguchi†, Katsuya Sato† and Shoichiro Fujisawa†

†Institute of Technology and Science, The University of Tokushima
2-1 Minami-josanjima, Tokushima, 770-8506 Japan
Email: {s.ito, 5007060396, katsuyas, s-fuji}@eco.tokushima-u.ac.jp

**Abstract**—In this paper, we introduce an EEG analysis technique to confirm an inter-individual difference in prefrontal cortex EEG with a single point sensing. The device for recording the EEG uses the dry-type sensor and a few numbers of electrodes. The EEG analysis adapts the feature mining on EEG pattern using a self-organizing map (SOM). The EEG patterns are determined based on the preference evaluation on sound listened to. In the preprocessing, we extract the EEG feature vector by calculating the time average on each frequency band which are $\theta$, low-$\alpha$, high-$\alpha$, low-$\beta$, high-$\beta$, respectively. To confirm the inter-individual difference, we do experiments using real EEG data. These results show that the learning results by SOM on each human are clearly different when using same initial weight values for the SOM.

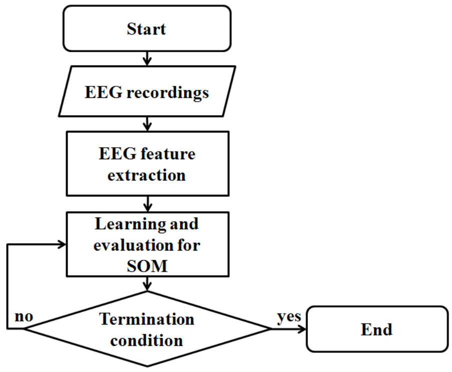


Figure 1: Procedure of the proposed method.

## 1. Introduction

It is an important issue that the electro cap with a large number of electrode is uncomfortable for human to wear and thus unsuitable for long-time recordings for using brain computer interface (BCI) in daily-life application. Therefore, we have attempted to construct the BCI using a compact device with dry-type electrodes. The target sensing point is the left lobe and single electrode is used. The EEG activities in the prefrontal pole have variability. The inter-individual difference is one of the factors in variability. Especially, the difference is of particular note when the sensing position is prefrontal cortex. However its reasons are not clearly. Therefore, this paper proposed a method to understand the inter-individual difference in EEG with single point sensing by analyzing the EEG.

Numerous approaches exist for analyzing EEG activity[1], such as analyzing the EEG features; power spectrum, spectral centroid, principal component analysis (PCA), independent component analysis (ICA), factor analysis (FA), *k*-nearest neighbor (*k*NN), linear discriminant analysis (LDA), neural network (NN), support vector machine (SVM), self-organizing map (SOM), etc. The SOM is capable of expressing the inter-individual difference by visualizing and classifying the EEG patterns. Because it is applied to confirm various multivariate data set, and has advantages over statistical and other non-traditional methods of cluster analysis.

In order to confirm the inter-individual difference, we perform experiments using a real EEG data.

## 2. Proposed methods

The proposed method consists of two phases; EEG recording and the measurement of the inter-individual difference in the EEG using a self-organizing map for the data mining. Fig. 1 shows the procedure of the proposed method.

### 2.1. EEG recordings

In EEG recording, we use the “MindTune (MT)” device, which was developed by TOSHIBA in Japan, to measure EEG activity. Generally, EEG systems use an electro cap. However, it is an important issue that the electro cap with a large number of electrode is uncomfortable for human to wear and thus unsuitable for long-time recordings for using BCI in daily-life application. Therefore, preparation of EEG recording before BCI operation takes a long time wearing the electro cap. Reducing the number of electrodes in BCI system is a critical issue. The MT uses the dry-type sensor and a few numbers of electrodes in the headphone. It does not need gel and/or water to wear the electrodes. Therefore we think it can alleviate uncomfortable feelings and can be used under realistic conditions. This methodology is a referential recording. The reference electrode is at the left ear and the exploring electrode is at Fp1 in the international 10-20 system.The EEG data obtained are sent to the computer every second through the serial port. The power spectra of EEG data per second are calculated by fast Fourier transform (FFT). The FFT data covers the frequency bands, $\delta$, $\theta$, low-$\alpha$, high-$\alpha$, low-$\beta$, high-$\beta$, low-$\gamma$ and high-$\gamma$.

After the EEG recording, the user completes an easy-questionnaire of preference evaluation on the sound listened to. This paper defines the EEG patterns based on results of the preference evaluation. The criteria of questionnaire is whether one likes the sound ("LikeSound"), dislikes it ("DislikeSound") or feels other ("Other"). The number of EEG patterns is three; LikeSound, DislikeSound and Other.

### 2.2. Measurement of inter-individual difference in EEG

To confirm the inter-individual difference in the EEG, we use the SOM in the data mining methods. The SOM is applied to confirm various multivariate data set, and has advantages over statistical and other non-traditional methods of cluster analysis.

The SOM is a means for automatically arranging statistical data so that alike input vectors are in general mapped close to each other. The resulting map avails itself readily to visualization. The distance relations between different data sets in the map are able to be illustrated in an intuitive manner. The SOM techniques have been successfully applied in a number of disciplines including speech recognition, image classification, document clustering and EEG analysis; especially using for the EEG feature visualization and the EEG pattern classification.

The algorithm for the inter-individual difference measurement is as follows:

1. The EEG feature vector is extracted for each EEG data patterns in the EEG data sets. First, the time series power spectra of five frequency bands that are $\theta$, low-$\alpha$, high-$\alpha$, low-$\beta$ and high-$\beta$ pick up in an EEG data pattern. Because the frequency bands of $\delta$, low-$\gamma$ and high-$\gamma$ all have special EEG meaning activity; they are not included in the EEG feature vector. Second, the sporadic rate of each frequency bands on each second during listening to the sound is calculated. Moreover, the discrete time average of the sporadic rate is computed. We regard as calculation results ($x(k), k = 1, 2, \ldots, 5$; 1: $\theta$, 2: low-$\alpha$, 3: high-$\alpha$, 4: low-$\beta$, 5: high-$\beta$) as the EEG feature vector. Finally, these operations are applied to all EEG data patterns.

2. The N-by-N map for the SOM sets including the nodes that consist of 5-dimensional vector as the weight vector.

3. The weight vectors are assigned randomly or 1.0 as an initialization.

4. The EEG data sets for learning are chosen based on the repeated random sub-sampling validation algorithm or all data sets. In the repeated random sub-sampling validation, 80 % in all data are chosen randomly as data sets for learning.

5. The weight vectors are updated recursively after the presentation of each input vector. As each input vector is presented, the Euclidian distance between the input vector and each weight vector is calculated using

$$D_{ij}(w_{ij}(k), x(k)) = \| x(k) - w_{ij}(k) \| . \quad (1)$$

   The winning node (denoted by subscript $c$) is specified by

$$d_c(k) \equiv min\ D_{ij}(k). \quad (2)$$

   The weight vectors are updated by

$$w'_{ij}(k) = w_{ij}(k) + \alpha * [x(k) - w_{ij}(k)] \ : \ i, j \in h_{ck}. \quad (3)$$

   where $\alpha$ means the learning rate factor, and $h_{ck}$ is the neighborhood function. The neighborhood function is typically a decreasing function of the distance on the two-dimension lattice between nodes $c$ and $k$. The standard neighborhood function is used.

$$h_{ck} = h_{ck}(0)(1 - LearNum/TotalLearNum) \quad (4)$$

   where $LearNum$ and $TotalLearNum$ indicate the number and total number of learning, respectively. The width $\sigma$ of the neighborhood function decreases during learning. The initial value ($h_{ck}(0)$) of the width for learning is half size of the map. This operation is repeated until the number of learning is met for more than a set number.

6. To evaluate, all learning EEG data patterns are mapped in the learned SOM. Then, the accuracy rate is computed based on the EEG patterns classification.

$$Accuracy = CorrectNumber/TotalNumber \quad (5)$$

   where the $CorrectNumber$ is the total number of correct answer by checking LikeSound, DislikeSound and Other. $TotalNumber$ means the total number of sounds listened to.

7. Operations 5 and 6 are repeated until being adapted to all chosen data sets.

8. Operations 3 to 7 are repeated until the number of trials is met for more than a set number.

## 3. Experiments

The subjects in this study comprised 5 persons: four males (average age 22.5 years old) and a female (age 22 years old). The experiment proceeded as follows: the EEG

Table 1: Kind of sounds listened to.

| | | | | |
|---|---|---|---|---|
| fire engine siren | wind bells sound | helicopter noise | cicada buzz | grade crossing |
| scotch tape | roar of waves | bush warbler buzz | mosquito | train noise |
| fireworks | soda water | unwrapping the paper | drill noise | frictional noise of styrene foams |

Table 2: Results of preference evaluation on sounds. Total indicates the total number of sounds on all and/or each subjects.

| | All | subject 1 | subject 2 | subject 3 | subject 4 | subject 5 |
|---|---|---|---|---|---|---|
| sex | - | male | male | male | male | female |
| LikeSound | 44 | 9 | 13 | 8 | 0 | 14 |
| DislikeSound | 178 | 35 | 48 | 39 | 24 | 32 |
| Other | 153 | 31 | 14 | 28 | 51 | 29 |
| Total | 375 | 75 | 75 | 75 | 75 | 75 |

device was positioned on the forehead of each subject; the subject then sat on a chair, closed his/her eyes, and remained quiet. The EEG was recorded more than once in the laboratory with environment noise during the experiment. The time table of each EEG recoding was 15 seconds (no sound) and 15 seconds (listening to a sound) as a set. After EEG recording, he/she completed the easy-questionnaire for preference evaluation on the sounds listened by checking LikeSound, DislikeSound and Other, respectively.

The total number of sounds listened to is 75 for each subject. Tables 1 and 2 show the kind of sounds listened to and the preference evaluation results on sounds listened to, respectively.

In the parameters for the SOM, the width of map and the learning rate $\alpha$ are 10-by-10 and 0.02, respectively. The number of learning and trials is 10,000 and 100, respectively. Figs. 2 and 3 show the results of the EEG feature map using the same initial weight values and the different initial weight values, respectively. Table 4 shows the results of the EEG pattern classification on learning data sets for the SOM.

## 4. Discussions

In Fig. 2, although the initial weight values on each subject were same, the contrast of the maps was not same. Especially, the variability of the contrast in (c) looked high. These results suggest that the EEG feature vectors on same EEG pattern in all EEG data sets may be not similar, and it was difficult to learn the all EEG data sets because of varying wide among the data sets of the same EEG patterns.

In Fig. 3, the contrast among the maps in each subject was similar compared with the results shown in Fig. 2. These results suggest that the EEG data sets in each subject are learned steadily.

Although the accuracy rate adapting data sets of all subjects was low, the results in each subject were high as shown in Table 4. These results suggest that the SOM cannot learn the EEG data patterns on all subjects because of the remarkableness of the inter-individual difference. Then, we confirmed that the SOM learned the EEG data patterns when adapting data sets on each subject, because the mean values and the standard deviation values of the accuracy rate were high and low, respectively.

From the results shown in Fig. 3 and Table 4, the inter-individual difference in EEG was clear using the SOM. We found that its difference was able to express the visualization results by the SOM after learning and the EEG pattern classification results.

## 5. Conclusions

We proposed a method to understand the inter-individual difference in EEG with single point sensing. The EEG device adapted the dry-type electrodes. The measurement position was left prefrontal pole (Fp1 in the 10-20 system). Moreover, the SOM was used for analyzing the EEG and visualizing the inter-individual difference.

In order to understand the inter-individual difference, we did experiments using real EEG data. The experimental results show that the learning results by SOM on each human were clear different when using same initial weight values for the SOM. Furthermore, the EEG pattern classification results on learning data sets for the SOM were better steadily.

Future work will involve effort to evaluate the EEG pattern classification on test data sets that are not learned by using bootstrap method and/or cross validation method in re-sampling techniques.

## References

[1] F. Lotte, M. Congedo, F. Lecuyer and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interface," *Journal of Neural Engineering*, 4, R1–R13 (2007).

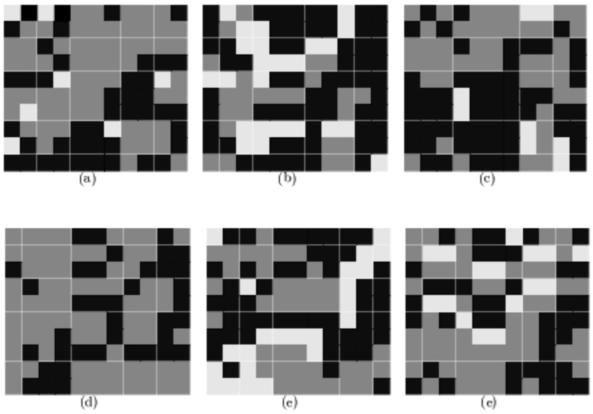

Figure 2: Maps on each subject when the initial weight values are same. (a) to (f) are subject 1 to 5 and all subjects, respectively. White, gray and black cells indicate LikeSound, Other, DislikeSound, respectively.

Table 3: Mean and S.D. of the accuracy rate for learning EEG data sets that are chosen randomly (100 trials).

| | all subjects | subject 1 | subject 2 | subject 3 | subject 4 | subject 5 |
|---|---|---|---|---|---|---|
| Accuracy | 0.61 ± 0.02 | 0.91 ± 0.03 | 0.87 ± 0.03 | 0.89 ± 0.03 | 0.91 ± 0.03 | 0.88 ± 0.03 |

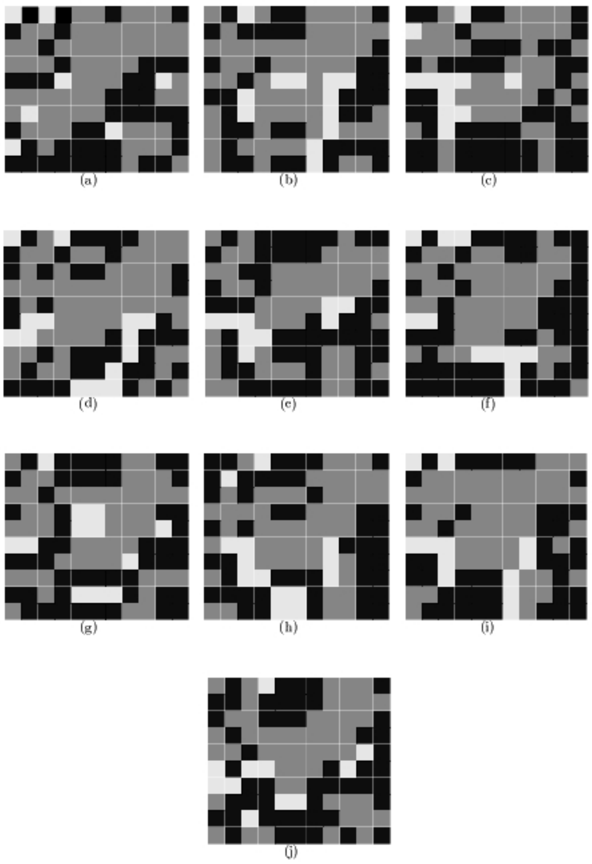

Figure 3: Example Maps on subject 1 the initial weight values are different. (a) to (j) are the iteration number 1 to 10, respectively.