Feature extraction method for EEG during motor imagery of limbs

Ryousuke Shibata, Kenji Hontani, Makoto Maeda and Katsuhiro Inoue
Kyushu Institute of Technology
680-4, Kawazu, Iizuka, Fukuoka 820-8502, Japan
E-mail: {shibata, hontani}@kiri.ces.kyutech.ac.jp, {inoue, maeda}@ces.kyutech.ac.jp

Abstract

Recently, many researchers have been studying Brain Computer Interface (BCI) by which we can operate objects with analyzing Electroencephalogram (EEG) instead of hands. One of them, by using event related synchronization from EEG during motor imagery of each hand, discriminated right or left hand's movement [1]. However, these methods enable us to discriminate only a few classes. Furthermore, they need a long-term training for improving their discrimination accuracy.

In this paper, we construct a system for discriminating motor imaged limbs using Movement Related Cortical Potentials (MRCPs), such as BereitschaftsPotential (BP), Negative Slope (NS) and Contingent Negative Variation (CNV), in which it does not need a long-term training.

1 Introduction

EEG is changed by activities, such as action, memory and attention. Brain Computer Interface (BCI) is to operate objects using such changes. These interfaces have been developed for people with disabilities, such as paralysis and movement disorder. The methods realizing them are roughly divided into 2 types: the one using stimulations of sounds or lights [2], and the other using motor imagination [3], [4]. In case of motor imagination, ERD/ERS (Event Related Desynchronization / Synchronization) is generally used. And a long-term training period is needed in order to obtain a satisfied determination rate. Although the stimulation method can discriminate among multi classes, it causes some strains that make subjects gaze on blinking lights. On the other hand, while the motor imagery method has no such strain, it can discriminate only a few classes [1].

In this paper, we construct an EEG discrimination system of motor imagery in 4 classes. Moreover, we target for enough discrimination rate without a long-term training. Hence we discussed 2 types of feature extraction methods using potential fluctuations of Movement Related Cortical Potentials (MRCPs) and barycentric positions of the source signal estimated by ICA [5], [6], [7].

2 Experimental Paradigm

2.1 Electrodes arrangement

Fig.1 shows electrode positions of the extended international 10-20 system stipulated by the international EEG Society. We measured a total of 15 channels at electrodes filled with blue color. They include C3, C4 and Cz which correspond to the motor area of the cerebral cortex. Besides them, we selected 2 electrodes on the outside of eyes for measuring eye movement and reference electrode respectively. In addition, eight channels added for measuring Electromyography (EMG) by placing two electrodes on each limb, in which we confirmed whether some actual movements had occurred during motor imagery.

Fig.1 Electrodes arrangement.

2.2 Experiment method

We carried out the following experiment to healthy 5 subjects whose average age was 22.4 years old. First, we had the subjects sit on a sofa in shielded room which was placed about 1m away from display. Second, the subjects heard a beep sound 2 sec after initial states as time chart shown in Fig.2. Then, the subjects imagined moving a limb corresponding to an instruction after 5 to 10 sec. In order to standardize the imagination of all subjects, they imagined the situation which is lifting up an object by each hand for the motor imagery of hands, and bending each ankle for the motor imagery of feet. The above operation for 10 seconds is...
referred to as 1 trial, and 40 trials were repeated in 1 session. An experiment consisting of 6 sessions was carried out in 1 day, and it is continued for 2 days. After EEG signals were recorded with sampling frequency 512Hz, a high pass filter over 0.1Hz was applied to the signals by built-in amplifier.

Fig. 2 Experiment time chart.

3 Method

3.1 A method using MRCPs
We investigated a method that using MRCPs. In this section, after confirming each component of MRCPs from the averaged waveforms, we explain a method to extract the features from single trial EEG data.

3.1.1 Pilot study
Although we evaluated some averaged waveforms at all the electrodes, we show only the waveforms that were obtained at C3 from subject 002. The waveforms for each class are shown in Fig.3.

Fig. 3 Averaged waveforms for Subject 002ms (at C3).

Fig.3 shows the different variations of EEG among limbs for subject 002ms. It has been confirmed that Negative-Slopes (NS) begin at about 0.5 sec before the motor imagery onset, in which they appear in period indicated as blue rectangle. Similarly, Contingent Negative Variations (CNV) begin at about 2 sec before the motor imagery onset, in which they appear in period indicated as red rectangle. From the figure, it has turned out that the amplitude of the waveform has a slight difference in both NS and CNV components between hand and leg.

3.1.2 Feature extraction in potential fluctuation
Based on the study on MRCPs, we tried to extract several slopes of the periods when NS and CNV components may appear. The extractions were performed from single trial EEG data. We define the slopes as shown in Fig.4 as features, where

\[ x_{ch}^1: \text{Slope of EEG } [-1.0, 0.0] \text{ sec}, \]
\[ x_{ch}^2: \text{Slope of EEG } [0.0, 1.0] \text{ sec}, \]
\[ x_{ch}^3: \text{Slope of EEG } [-1.5, -1.0] \text{ sec}. \]

Fig. 4 Each slope extracted from single trial EEG data.

3.2 A method using barycentric position
For more practical BCI system, we should use inexpensive instruments. However, they do not have enough electrodes to measure center of the head. Thus, we use ICA for estimating barycentric position of the source signal as a pre-process in order to achieve BCI without center electrode.

3.2.1 ICA method
ICA is a method for separating signals. Suppose observed signals \( x \) are expressed as a linear combination of original signals \( s \) which are statistically independent. Then we can obtain separated signals \( y \) corresponding to the original signals \( s \) by estimating confusion matrix \( A \). Fig.5 shows a process in which observed signals \( x \) are generated from original signals \( s \). By using the separated matrix \( W \) and confusion matrix \( A \), these signals are obtained as follows.

\[ x = As, \quad y = Wx. \]

Fig.5 ICA process.
In brief, ICA can estimate \( y \) so that \( W \) may become the same as the inverse of \( A \), under the assumption that the original signals \( s \) are statistically independent.

### 3.2.2 Selection of criteria

ICA has a problem that the order of \( y \) is different from \( s \). Therefore, it is unknown which signal in separated signals is related to motor imagery. In this paper, we set 4 assumptions of target source signal related to motor imagery.

**Assumption 1:**
Target source signal is the most activity signal. Then, the signals may have the largest amplitude.

**Assumption 2:**
Target source signal is the most activity signal. Then, the signals may have the largest absolute value.

**Assumption 3:**
Target source signal have a peak signal (event related potential) related to motor imagery and the peak may appear at just after the instruction cue appeared.

**Assumption 4:**
Target source signal have a peak signal (event related potential) related to motor imagery and the peak may appear at just before the motor imagination starts.

Therefore, in this paper, the following 4 criteria were adopted according to the above mentioned assumptions.

**Criterion 1:**
signal having the largest amplitude,

**Criterion 2:**
signal having the largest absolute value,

**Criterion 3:**
signal whose peak value comes the most quickly,

**Criterion 4:**
signal whose peak value comes the most slowly.

Fig. 6 shows examples of choosing method of target source signal based on each criterion.

### 3.2.3 Estimation of barycentric position

We consider the normalized inverse of \( W \) and \( A \) almost equivalent. Suppose that \( s \) is distance attenuation, then the barycentric position can be computed by the parameters of \( A \) and the 3D position \( x \) of each electrode.

\[
W' = \frac{W^{-1}}{\|W\|} = \begin{bmatrix}
W'_{11} & \cdots & W'_{1N} \\
\vdots & \ddots & \vdots \\
W'_{N1} & \cdots & W'_{NN}
\end{bmatrix}
\]

\[
u(g_x, g_y, g_z) = \begin{bmatrix}
u_1(g_x, g_y, g_z) \\
\vdots \\
u_N(g_x, g_y, g_z)
\end{bmatrix} = \begin{bmatrix}
w_1^{11} & \cdots & w_1^{1N} \\
\vdots & \ddots & \vdots \\
w_N^{11} & \cdots & w_N^{1N}
\end{bmatrix}^T \begin{bmatrix}
x_1(g_x, g_y, g_z) \\
\vdots \\
x_N(g_x, g_y, g_z)
\end{bmatrix}
\]

The estimated barycentric positions are mapped to a head model. We use the 3D positions of barycentric position as features.

### 3.3 Discrimination method

Each feature mentioned in Section 3.1.2 and 3.2.3 is employed to Bayes discrimination method as the following formulas. The discrimination rates are evaluated by using 10-fold-crossvalidation.

\[
k_{ch}^* = \text{ArgMin}_{k=1}^{5} \left\{ -\frac{1}{2} (x_{ch} - M_{ch}^k)^T \Sigma_{ch}^{-1} (x_{ch} - M_{ch}^k) + \frac{1}{2} \ln |\Sigma_{ch}| - \ln \Pr(C_{ch}) \right\} \tag{1}
\]

\[
M_{ch}^k = \frac{1}{N} \sum_{i=1}^{N} x_{ch(i)}^k, \tag{2}
\]

\[
\Sigma_{ch}^k = \frac{1}{N-1} \sum_{i=1}^{N} (x_{ch(i)}^k - M_{ch}^k)(x_{ch(i)}^k - M_{ch}^k)^T, \tag{3}
\]

where \( x_{ch} = [x_{ch1}^2, x_{ch2}^2, x_{ch3}^2]^T \), ch: electrode.

### 4 Experimental results

#### 4.1 Discrimination result based on MRCPs

##### 4.1.1 Result of using MRCPs

We tried 2 approaches using features mentioned in Section 3.1.2.

First, we extracted the 3 slopes at the electrodes C3, C4 and Cz, so that we constructed the 9-dimensional feature vector. We refer to the vector as an augmented feature vector. Hence, the first approach is to apply Bayes discrimination method using the augmented feature vector \( x_{ch} = [x_{ch1}^2, x_{ch2}^2, x_{ch3}^2, x_{ch1}^2, x_{ch2}^2, x_{ch3}^2, x_{ch1}^2, x_{ch2}^2, x_{ch3}^2]^T \).

Second, we used a majority rule. After the discrimination...
candidates were independently determined at three electrodes, the final decision was determined by three candidates’ votes. If the vote was not able to determine one class, we discarded the sample.

In order to compare with the above approaches, we investigated which electrode had the highest discrimination rate in all the electrodes. We refer to the discrimination rate as \( \text{Max}_\text{ch} \).

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\text{subject ID} & \text{Max}_\text{ch} & \text{Augmented feature vector} & \text{Majority rule} \\
\hline
\text{day1} & 002ms & 003ws & 004ny & 005nk & 006kt \\
\text{day2} & 002ms & 003ws & 004ny & 005nk & 006kt \\
\hline
\end{array}
\]

It has confirmed from Fig.7 that 3 subjects, 002ms, 003ws and 006kt, achieve much higher discrimination rate than chance probability of 25%. They are subjects who MRCPs components have appeared in their averaged waveforms. Also the discrimination rates using Augmented feature vector and Majority rule are better than \( \text{Max}_\text{ch} \). While using Augmented feature vector increases the discrimination rate 5% to 10%, using Majority rule increases 10% to 15%. However, the number of discarded data by Majority rule reaches about 50% in all subjects. Therefore, Augment feature vector seems to be better feature in real BCI system.

4.1.2 Learning effect of subject based on MRCPs

In order to investigate the learning effect, the data measured among 1~N-1 sessions were used as training data, and the data measured in N session were used as test data. We show the results of the discrimination rate in Fig.8. The discrimination was carried out discriminate using augmented feature vector method in session 3.3.

From this result, sessions 1-3 were the result is low discrimination rate with all subjects. We speculate that this is not enough number of training data, because the covariance matrix can not be correctly estimated. However, discrimination rate is improved by increasing the number of data from session 1-3. Then, it was found that to obtain a subsequent stable discrimination rate. From above mentioned result, the number of four sessions data may be sufficient to create a feature space in actual fields. In addition, if the data is to focus on the data after the session 7 that was a different date training data, these discrimination rates did not change significantly. Actually, it shows the result of obtaining the discrimination rate that of the first day of data set the training data, second day set the test data.

In our studies in which AR-model and Bayes decision rule was used to discriminate between right hand imagery and left hand imagery, about 10 days was needed to obtain high accuracy. But these results suggest that learning period can be reduced by using the proposed method.

4.2 Discrimination result based on ICA

4.2.1 Result of ICA (4 class)

Fig.9 shows an example of the estimated barycentric positions of subject 002ms, in which we chose the signal related to motor imagery using Criterion 1.

The black dots represent the electrode positions and the blue dots represent barycentric positions. As for all the results, there is almost no difference in the height among barycentric positions for each limb. Note that the barycentric position estimated when each leg was imaged appeared in contralateral side. For example, the barycentric positions appeared right region in case of left leg imagination.

Fig.10 shows the results of discrimination rate for each criterion. 2 subjects achieved about 40% discrimination rate by using criterion 1 and 2. The rates were lower in criterion 3,4 because left and right position of biases were not confirmed for all subjects.

4.2.2 Result of ICA (2 class)

In previous section, we researched 4 classes discrimination. However, it was not so easy to get available discrimination rate in 4 classes. Therefore, we deal with the 2 classes problem and discuss which criterion is the most effective. Fig.11 shows each result of criterion. In these figures, “Right hand” omitted as “RH”.

It has confirmed from Fig.11 (a) that 3 subjects have the limbs combination that achieves 65% or more. They have higher discrimination rate than “Right Hand-Left Foot” and “Left Hand-Right Foot” in common. Next, from (b), 2 subjects have the combination that achieves 65% or more. The combination similar to the criterion1. Finally, it has confirmed from (c) and (d) that the discrimination rate is low overall. The subjects who have achieved 65% or more at Criterion1 and 2 are declining. These results suggest that more effective criterion is amplitude rather than peak time.
Fig.9 Barycentric positions estimated based on criterion1 (002 ms).
Fig.10 Discrimination rate in 4 classes based on each criterion.

Fig.11 Discrimination rate based on each criterion.

For Fig.12 (004ny), this subject has low discrimination rate in each method since he does not have features such as CNV and NS by each limb.

5 Conclusions

In this paper, two types of features were investigated in order to develop the effective feature extraction methods for EEG during motor imagery of limbs. One was feature values based on MRCPs and another was feature values based on the barycentric position of the source signal estimated by ICA.

It was confirmed that 60% - 80% discrimination rates in 4 classes using single trial data were obtained by the method based on MRCPs. And it was also confirmed that the time required for learning was shortened by this method compared with the conventional method based on ERD/ERS.

On the other hand, for the method based on the position of barycenter, the difference tendency was confirmed although the satisfactory determination rate was not obtained. This result shows that more essential information which is related to motor imagery may be extracted by this method by improving the selection method of the source signal.

In this study, the method based on MRCPs with augmented feature vector was the best method and the applicability of it to BCI was confirmed. But, this method has the demerits that the beep sound is needed as the previous notice stimulation that directs the beginning of imagination, while such beep sound is not needed on the conventional method which is based on ERD/ERS. We think that the optimal method should be developed by combining the information of MRCPs and ERD/ERS and the position of barycenter of estimated source signal. These studies are future works.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 25282177.
References


