A Bayes Classifier Considering Environmental Change for Multivariate Signal Data

1st Itaru Aso
Dept. Information Systems Engineering
Toyama Prefectural University
Kurokawa, Imizu, Toyama, Japan
t855001@st.pu-toyama.ac.jp

2nd Mamoru Ota
Dept. Information Systems Engineering
Toyama Prefectural University
Kurokawa, Imizu, Toyama, Japan
ota@pu-toyama.ac.jp

3rd Koji Okuhara
Dept. Information Systems Engineering
Toyama Prefectural University
Kurokawa, Imizu, Toyama, Japan
okuhara@pu-toyama.ac.jp

***Abstract*—In this paper, we suggest learning algorithm of a high precision classifier for multivariate signal. The method deals with environmental influences. In this proposal technique, we define the features of the classification target and the environment as population parameters of probability distribution. We estimate the parameters by using the Bayesian inference. The Bayesian decision rule is used for the selection of similar environment properly in the proposed method. We try to evaluate the influence of the environmental change. In the numerical experiments, we verify that the proposed method has high classification accuracy. As the results, we show that our method can adapt environmental influence.**

***Index Terms*—signal processing, Bayesian inference, EEG**

I. INTRODUCTION

We can accumulate a large amount of data of various types because measuring techniques are developing nowadays. The demand for techniques that analyzes acquired data comprehensively is increasing, and application examples of machine learning are also increasing.

In particular, there are many researches of machine learning. The aim of these researches are identification and prediction some states. The techniques apply to signal data of various fields such as finance, audio and biological.

As one example, there is a research that learns the relation between the characteristics of sound and emotions. The research classified feelings from voice [1]. This research estimates feature of measurement voice obtained by the principal component analysis (PCA) and the factor analysis, and the relationship between audio data and feeling was found by the features.

In addition, types of musical instrument (e.g. piano and guitar) from audio signal was classified to model a change of tone efficiency [2]. The instrument classification by using Bagging was evaluated. 10 types of actions of exercise to radio by using three dimensional time series data were classified to identify the separation between actions [3]. A research of action classification used PCA and Linear Discriminant Analysis (LDA). It classifies the 10 types of actions by using the Bayesian theory.

In the medical field of applied researches, the platelet data was used as input, and was identified patients with hepatitis B and patients with hepatitis C. There is another paper that classifies the action of rock-paper-scissors by Mahalanobis distance [4,5]. In this research, one point electroencephalogram (EEG) in measurement was used.

Such methods will be used for more expanded application fields in the near feature. It is therefore important to establish the machine learning technique.

Considering diversification of data measurement, it becomes necessary to deal with environmental influence. In other words, the accuracy of machine learning may decrease when we obtain data from measuring instruments under external factors (e.g. weather and physical condition).

Carrying out measurements not affected by the environmental influences is better for improve the quality of data; however, it is almost impossible because there are many restrictions on the measurement. The robust analysis in consideration environmental influence is therefore needed for high accuracy classifier. The method can solve a classification problem considering environmental change.

In the present study, we proposed the classification method with high accuracy. Our proposed method is based on Bayesian inference. The aim of the method is to adapt environmental influences. We defined the features of the classification target and the environment as population parameters of probability distribution. There are more examples of studies using Bayesian inference such as estimated shopping mode choice and driving behavior intention using information of driver's line sight, and estimating fatigue life of instruments [6,7,8].

Our proposed method can choice the optimum environment so this method can also deal with environmental influence. In numerical experiments, we evaluated the performance of the proposed method by using EEG data. An analysis of adapting environmental influences for the multivariate signal was performed.

II. PROPOSAL OF STATE IDENTIFICATION METHOD

D-dimensional vector datasets were obtained from measuring instruments, with sample number was N. The number of measurements was C. The method classify K states considering environmental influence as follows:

A. Problem setting

X is an observed dataset, and it is acquired from measuring instruments. X is defined as

$$\boldsymbol{X} = \left\{ \boldsymbol{X}^{(\mathrm{c})} \,\middle|\, c = 0, 1, \ldots, C-1 \right\}, \tag{1}$$

where $\boldsymbol{X}^{(\mathrm{c})}$ is defined by

$$\boldsymbol{X}^{(\mathrm{c})} = \left\{ \boldsymbol{x}_n^{(\mathrm{c})} \in \mathbb{R}^D \,\middle|\, n = 0, 1, \ldots, N-1 \right\}, \tag{2}$$

and c represents number of measuring instruments. Also, teacher label of environment is given by

$$\boldsymbol{E} = \{\, \boldsymbol{e}_n \in \mathbb{T}_M \mid n = 0, 1, \ldots, N-1 \,\}, \tag{3}$$

where M is the number of environmental types. Teacher labels of states is given by

$$\boldsymbol{Z} = \{\, \boldsymbol{z}_n \in \mathbb{T}_K \mid n = 0, 1, \ldots, N-1 \,\}. \tag{4}$$

Teacher label represents K states of identification target. It is assumed that each $\mathbb{T}_L$ is given by

$$\mathbb{T}_L \equiv \left\{ \boldsymbol{z} \in \{\, 0, 1 \,\}^L \,\middle|\, \sum_{\ell=0}^{L-1} [\boldsymbol{z}]_\ell = 1 \right\}. \tag{5}$$

$\boldsymbol{X}_{(\boldsymbol{e},\boldsymbol{z})}$ is defined by

$$\boldsymbol{X}_{(\boldsymbol{e},\boldsymbol{z})} = \left\{ \boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \,\middle|\, c = 0, 1, \ldots, C-1 \right\}, \tag{6}$$

where $\boldsymbol{e}$ is the measurement environment type and $\boldsymbol{z}$ is the identification target state. We also defined $\boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ as

$$\boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \equiv \left\{ \boldsymbol{x}_n^{(\mathrm{c})} \in \mathbb{R}^D \,\middle|\, (\boldsymbol{e}_n, \boldsymbol{z}_n) = (\boldsymbol{e}, \boldsymbol{z}), n = 0, 1, \ldots, N-1 \right\}, \tag{7}$$

where each $\boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ is the dataset with features of specified environment and state $(\boldsymbol{e}_n, \boldsymbol{z}_n) = (\boldsymbol{e}, \boldsymbol{z})$.

The proposed method uses the Bayesian inference to learn features of each $\boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$. We make an assumption the probability distribution following the multivariate Gaussian distribution. The mean vector and the variance-covariance matrix are estimated [9, 10].

B. Learning rule based on the Bayesian inference

Under the assumption mentioned above, prior probability $P_0(\boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})})$ is defined by

$$\begin{aligned} P_0(\boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}) &= \mathcal{N}_D(\boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} | \boldsymbol{m}_0, (\beta_0 \boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})})^{-1}) \\ &\quad \times \mathcal{W}_D(\boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} | \alpha \boldsymbol{I}_D, \nu_0). \end{aligned} \tag{8}$$

$\boldsymbol{\mu}_{(\boldsymbol{e},\boldsymbol{z})}$ and $\boldsymbol{\Lambda}_{(\boldsymbol{e},\boldsymbol{z})}$ denote mean vector and precision matrix for the data set $\boldsymbol{X}_{(\boldsymbol{e},\boldsymbol{z})}$, respectively. Posterior probability $Q(\boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} | \boldsymbol{X}^{(\mathrm{c})})$ is also defined by

$$\begin{aligned} &Q(\boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} | \boldsymbol{X}^{(\mathrm{c})}) \\ &= \mathcal{N}_D(\boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} | \boldsymbol{m}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, (\beta^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})})^{-1}) \\ &\times \mathcal{W}_D(\boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} | \boldsymbol{W}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \nu^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}), \end{aligned} \tag{9}$$

where $\mathcal{N}_D(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1})$ expresses probability density function of D dimensional Gaussian distribution with mean vector μ and precision matrix $\boldsymbol{\Lambda}$. Also, $\mathcal{W}_D(\boldsymbol{\Lambda}|\boldsymbol{W}, \nu)$ denotes probability density function of D-dimensional Wishart distribution, and $\boldsymbol{I}_D$ is the D-dimensional identity matrix.

The hyper parameters of posterior probability are calculated by

$$\boldsymbol{m}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} = \frac{\beta_0 \boldsymbol{m}_0 + N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \bar{\boldsymbol{x}}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}}{\beta_0 + N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}}, \tag{10}$$

$$\beta^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} = \beta_0 + N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \tag{11}$$

$$\begin{aligned} (\boldsymbol{W}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})})^{-1} &= \alpha^{-1} \boldsymbol{I}_D + N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \boldsymbol{S}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} + \frac{\beta_0 N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}}{\beta_0 + N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}} \\ &\quad \times (\bar{\boldsymbol{x}}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} - \boldsymbol{m}_0)(\bar{\boldsymbol{x}}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} - \boldsymbol{m}_0)^{\mathrm{T}}, \end{aligned} \tag{12}$$

$$\nu^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} = \nu_0 + N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}. \tag{13}$$

Then, $N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$, $\bar{\boldsymbol{x}}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ and $\boldsymbol{S}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ are given by

$$N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \equiv \left| \boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \right|, \tag{14}$$

$$\bar{\boldsymbol{x}}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \equiv \frac{1}{\left| \boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \right|} \sum_{n=0}^{N-1} \mathbb{I}[(\boldsymbol{e}_n, \boldsymbol{z}_n) = (\boldsymbol{e}, \boldsymbol{z})] \boldsymbol{x}_n^{(\mathrm{c})}, \tag{15}$$

$$\begin{aligned} \boldsymbol{S}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} &\equiv \frac{1}{\left| \boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \right|} \sum_{n=0}^{N-1} \mathbb{I}[(\boldsymbol{e}_n, \boldsymbol{z}_n) = (\boldsymbol{e}, \boldsymbol{z})] \boldsymbol{x}_n^{(\mathrm{c})} \left(\boldsymbol{x}_n^{(\mathrm{c})} \right)^{\mathrm{T}} \\ &\quad - \bar{\boldsymbol{x}}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \left(\bar{\boldsymbol{x}}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \right)^{\mathrm{T}}, \end{aligned} \tag{16}$$

where $N^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ is sample number of $\boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$. Also, $\bar{\boldsymbol{x}}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ denotes the sample mean vector of $\boldsymbol{x}_n^{(\mathrm{c})}$ and $\boldsymbol{S}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ expresses sample variance-covariance matrix of $\boldsymbol{x}_n^{(\mathrm{c})}$.

$\mathbb{I}[(\boldsymbol{e}_n, \boldsymbol{z}_n)$ is defined by

$$\mathbb{I}[(\boldsymbol{e}_n, \boldsymbol{z}_n) = (\boldsymbol{e}, \boldsymbol{z})] \equiv \begin{cases} 1 & (\boldsymbol{e}_n, \boldsymbol{z}_n) = (\boldsymbol{e}, \boldsymbol{z}) \\ 0 & (\boldsymbol{e}_n, \boldsymbol{z}_n) \neq (\boldsymbol{e}, \boldsymbol{z}). \end{cases} \tag{17}$$

If measurements data $\{\, \boldsymbol{x}^{(\mathrm{c})} \,\}$ is given in advance, we can obtain acceptance probability of the environment $\boldsymbol{e}$ and state of identification target $\boldsymbol{z}$. $P\left(\boldsymbol{e}, \boldsymbol{z} | \boldsymbol{X}, \{\, \boldsymbol{x}^{(\mathrm{c})} \,\}\right)$ is defined by

$$\begin{aligned} &P\left(\boldsymbol{e}, \boldsymbol{z} | \boldsymbol{X}, \left\{\, \boldsymbol{x}^{(\mathrm{c})} \,\right\}\right) \\ &\propto \prod_{c=0}^{C-1} \int\!\!\int \mathcal{N}_D(\boldsymbol{x}^{(\mathrm{c})} | \boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \left(\boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \right)^{-1}) \\ &\times Q(\boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} | \boldsymbol{X}^{(\mathrm{c})}) \mathrm{d}\boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \mathrm{d}\boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} \\ &= S_D(\boldsymbol{x}^{(\mathrm{c})} | \boldsymbol{m}^*, \boldsymbol{Q}^*, f^*). \end{aligned} \tag{18}$$

Then $\boldsymbol{m}^*, \boldsymbol{Q}^*, f^*$ are given by

$$\boldsymbol{m}^* = \boldsymbol{m}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \tag{19}$$

$$\boldsymbol{Q}^* = (\nu^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} - D + 1)\frac{\beta^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}}{\beta^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} + 1}\boldsymbol{W}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \tag{20}$$

$$f^* = \nu^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})} - D + 1. \tag{21}$$

S_D is the multivariate t-distribution with degree of freedom f, and it is defined by

$$\begin{aligned} &S_D(\boldsymbol{x}|\boldsymbol{m}, \boldsymbol{Q}, f) \\ &= \frac{\Gamma(\frac{f}{2}+\frac{D}{2})}{\Gamma(\frac{f}{2})}|\boldsymbol{Q}|^{\frac{1}{2}}\left(1+\frac{1}{f}(\boldsymbol{x}-\boldsymbol{m})^T\boldsymbol{Q}(\boldsymbol{x}-\boldsymbol{m})\right)^{-\frac{f}{2}-\frac{D}{2}}, \end{aligned} \tag{22}$$

where $\Gamma(\cdot)$ is the Gamma function. $P\left(\boldsymbol{e}, \boldsymbol{z}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}\right)$ must be standardized to

$$\sum_{(\boldsymbol{e},\boldsymbol{z})\in\mathbb{T}_M\times\mathbb{T}_K} P\left(\boldsymbol{e}, \boldsymbol{z}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}\right) = 1. \tag{23}$$

If M becomes large, the measuring environment of $\left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}$ can be found the similar environment within measuring environmental types of learning data. The measuring environment $\boldsymbol{e}$ of $\left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}$ is treated as the element of $\mathbb{T}_M$.

As a simple example for classification states, we can consider the method that the classification result equals to environmental state with the maximize acceptance probability.

Then, $\boldsymbol{e}^*_1$ and $\boldsymbol{z}^*_1$ are given by

$$(\boldsymbol{e}^*_1, \boldsymbol{z}^*_1) = \underset{\boldsymbol{e},\boldsymbol{z}\in\mathbb{T}_M\times\mathbb{T}_K}{\arg\max}\ P\left(\boldsymbol{e}, \boldsymbol{z}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}\right), \tag{24}$$

where $\boldsymbol{z}^*_1$ denotes the result of state classification and $\boldsymbol{e}^*_1$ denotes the optimum environment.

When the calibration data $\left\{ \boldsymbol{x}^{(\mathrm{c})}_{\mathrm{cal}} \right\}$ and state of target classification $\boldsymbol{z}_{\mathrm{cal}}$ are known, we can identify the measuring environment by using its information. Then, the optimal environment $\boldsymbol{e}^*_2$ with the maximum a posteriori probability is given by

$$\begin{aligned} \boldsymbol{e}^*_2 &= \underset{\boldsymbol{e}\in\mathbb{T}_M}{\arg\max}\, P\left(\boldsymbol{e}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})}_{\mathrm{cal}} \right\}, \boldsymbol{z}_{\mathrm{cal}}\right) \\ &= \underset{\boldsymbol{e}\in\mathbb{T}_M}{\arg\max}\, P\left(\boldsymbol{e}, \boldsymbol{z}_{\mathrm{cal}}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})}_{\mathrm{cal}} \right\}\right). \end{aligned} \tag{25}$$

Using the optimal environment estimated from the calibration data, we can classify the states from the continuous measuring data $\left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}$.

When $\boldsymbol{e}^*$ is obtained by Eq. (26), $\boldsymbol{z}^*_2$ is given by

$$\boldsymbol{z}^*_2 = \underset{\boldsymbol{z}\in\mathbb{T}_K}{\arg\max}\, P\left(\boldsymbol{e}^*, \boldsymbol{z}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}\right). \tag{26}$$

This classification method is expected to perform stable and highly accurate classification against changes of measuring environment compared to Eq. (24).

III. Numerical experiments

A. Outline

In this study, we used ULTRACORTEX MARK 4 as a headset and a Cyton board manufacured by OpenBCI[1)]. By using the equipment, we can obtain EEG data. We evaluated performance of proposed method to apply classification problems. In this problems, 4 states (no action and rock-paper-scissors classification) are considered.

In experiments, we used EEG signal data. Data of measuring positions can be expressed as multidimensional data. EEG data may not always be able to acquire data of the same characteristics due to the feature change depend on subject physical condition and shifting measuring position. The prediction accuracy can be decrease for this reason.

The number of subjects are 4 people (subject A, subject B, subject C, subject D).

We suggested the robust analysis method that deal with such environmental influence. Our method is based on theory of Section 2.

Obtaining the biological data is large burden for people. The amount of the obtained data are limited. It is therefore necessary to classify states from small amount of acquired data because it is difficult to acquire the enormous data.

Through the experiments, we show that the method has the effectiveness to be able to adapt environmental influence.

B. Experiment using EEG data

In experiments, a subject attached the EEG performed rock-paper-scissors action following instructions of the sequence (refer to Fig. 1). All slides are displayed for 1 second. The subject keeps standby state during that “3” and “2” of slides are displayed. In the third slide, rock-paper-scissors of the images are displayed.

A subject makes an action so as to win the rock-paper-scissors hand when the white slide is displayed.

Measuring positions of EEG obeyed the ten-twenty electrode system (refer to Fig. 2). We acquired EEG data from 8 measurement positions (Fp1, Fp2, C3, C4, P7, P8, O1, O2).

The experiment was carried out during 4 times because it is considered that the environmental influence depend on physical condition of a subject and shifting measurement positions.

We defined a set data (4 seconds) from the standby state to the rock-paper-scissors action states. We obtained 100 sets of EEG data in a day. We acquired 400 sets of the EEG data during for 4 times.

C. Application for the proposed method to EEG data

There are two ways to analyze EEG data. One way takes the reference points of potential from earlobes. The other way is that the average value of each measuring positions is regarded as the reference point. In this study, we adopted the latter way. The way is described below.

[1)]http://openbci.com

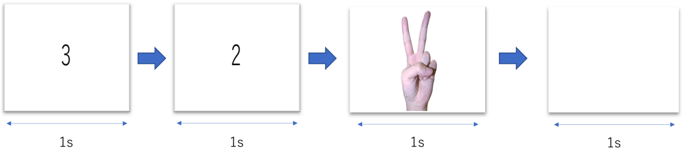

Fig. 1. Flow of created slide

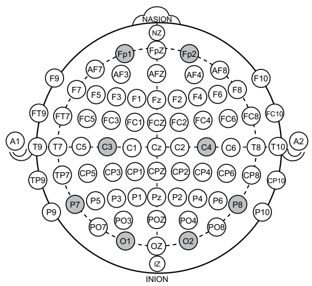

Fig. 2. Ten-twenty electrode system

If $\boldsymbol{S}(t) = [S^{(1)}(t), S^{(2)}(t), \cdots S^{(C)}]^{\mathrm{T}}$ is the obtained EEG data, the mean signal $\boldsymbol{\mu}(t)$ is obtained by

$$\boldsymbol{\mu}(t) = \frac{1}{C}\sum_{c=1}^{C} \boldsymbol{S}^{(c)}(t), \tag{27}$$

where C is the number of the measuring positions.

At the measurement position c, $\boldsymbol{S}_{\mathbf{0}}^{(c)}(t)$ is given by

$$\boldsymbol{S}_{\mathbf{0}}^{(c)}(t) = \boldsymbol{S}^{(c)}(t) - \boldsymbol{\mu}(t)\mathbf{1}^{\mathrm{T}}. \tag{28}$$

We obtained power spectra of acquired data by using fast Fourier transform (FFT) with Hamming window. The frequency band used for the analysis was 3-50 Hz. We divided the frequency band into 12 bands. The average powers of the sub-bands were used as the feature quantity for classification.

Also, using the hyper parameters of the prior probability are set to be $\beta_0 = 0.1$, $\boldsymbol{m}_{\mathbf{0}} = \mathbf{0}$, $\nu = 12$, $D = 12$, $K = 3, 4$.

D. Environmental influence

In order to check the existence of environmental influence, environmental data of every subjects were applied the frequency analysis in each measuring sites. The environmental data are regarded as EEG data with the standby state. As the result of the analysis, features of EEG data every observation was different. This difference was seen the observation of every subjects. It suggested that there was an influence by the physical condition of subject and deviation from measuring points. In particular, Fig. 3 shows average power spectra of 100 sets in the O2. The position shows a clear difference.

IV. The Result of Numerical Experiments

To classify the rock-paper-scissors actions, we used 1 second data while instruction slide is displayed. The 2 seconds data of the standby state are used as the environment data. The 3 states classification has rock-paper-scissors classes. The 4 states classification has rock-paper-scissors classes and the environmental class.

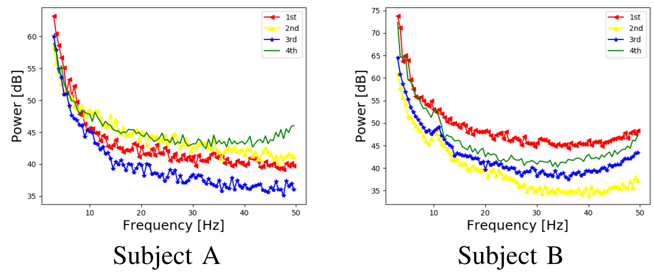

Fig. 3. Graph of frequency analysis of O2

We performed 6 types of verification. First, we executed 3 states classification and 4 states classification when we can ignore the environmental influence. Next, we performed two types of 3 states classifications that were adapting environmental influence and not adapting environmental influence. Finally, we also enforced two types of 4 states classifications that were adapting environmental influence and not adapting environmental influence.

As shown in Table I and Table II, the number of the training data and the test data are used the 6 types of verification. The first 60 sets of acquired data were used as training data. Also, the remaining 40 sets were used as test data.

In 4 states classification, the training data were 60 sets data of rock-paper-scissors classes and 60 sets data of environmental class. Also, the test data were 40 sets data of rock-paper-scissors classes day and 40 sets data of environmental class.

In Table II, test data of rock-paper-scissors classes for 3 times in not adapting environmental influence shows the number of 120. Also, test data of rock-paper-scissors classes and environment class for 3 times denotes the number of 240.

When we execute the classification adapting environmental influence, we obtain the classification result by using eq. (27) after finding the optimum environment of train data based on eq. (26).

TABLE I
The Number of the Train Dataset and the Test Dataset with Continuous Observation Data

3 class	training data	60
	test data	40
4 class	training data	120
	test data	80

A. 3 states classification without environmental influence

Table III shows the result of the classification without environmental influence.

In this study, changing parameter is only α. The values of α and the classification accuracy of the training data are shown in Fig. 4. The values of α and classification accuracy of the test data are shown in Fig. 5. Subject A obtained the highest accuracy of 53 % in 2nd.

TABLE II
THE NUMBER OF THE TRAIN DATASET AND THE TEST DATASET WITH ENVIRONMENTAL INFLUENCE

		Correspondence	No Correspondence
3 class	training data	60	60
	test data	40 1)	120 2)
4 class	training data	120	80
	test data	80 1)	240 2)

1) The optimum environment
2) For 3days

TABLE III
THE ACCURACY OF 3 STATES CLASSIFICATION WITH CONTINUOUS OBSERVATION DATA

	1st	2nd	3rd	4th
α	0.000794	0.000114	0.000154	0.000384
train(subject A)	0.81	0.51	0.66	0.55
test(subject A)	0.38	0.53	0.43	0.41
train(subject B)	0.61	0.51	0.66	0.55
test(subject B)	0.50	0.45	0.47	0.45
train(subject C)	0.58	0.65	0.55	0.60
test(subject C)	0.40	0.37	0.43	0.45
train(subject D)	0.63	0.60	0.53	0.67
test(subject D)	0.40	0.38	0.45	0.43

Since the accuracy of the test data was relatively low and the classification accuracy of the training data was high, we presume that the overfitting occurred in 1st. On the other hand, because the accuracy of the test data and the accuracy of the train data did not change in the other three days, we did not consider causing overfitting.

B. 4 states classification with continuous observation data

Table IV shows the result of 4 states classification without environmental influences.

Fig.6 shows the value of α and classification accuracy graphs of the training data. Fig. 7 shows the value of α and classification accuracy graphs of the test data. Subject D observed the highest accuracy of 56 % in 4th.

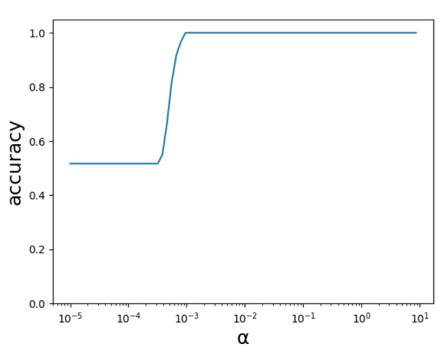

Fig. 4. The values of parameter α and the accuracy of 3 states classification in training data

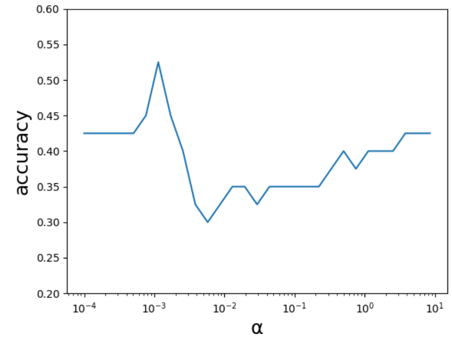

Fig. 5. The values of parameter α and the accuracy of 3 states classification in test data

TABLE IV
THE ACCURACY OF 4 STATES CLASSIFICATION WITH CONTINUOUS OBSERVATION DATA

	1st	2nd	3rd	4th
α	0.0195	0.0046	0.0017	0.0011
train(subject A)	0.55	0.75	0.52	0.52
test(subject A)	0.52	0.53	0.55	0.50
train(subject B)	0.68	0.60	0.63	0.58
test(subject B)	0.48	0.55	0.53	0.50
train(subject C)	0.62	0.56	0.65	0.60
test(subject C)	0.48	0.50	0.49	0.53
train(subject D)	0.65	0.70	0.62	0.58
test(subject D)	0.49	0.45	0.56	0.53

C. 3 states classification with environmental influence

We verified two classifications that were not adapting environmental influence and adapting environmental influence.

If the data for one day was the calibration data, we selected the optimum the environment from the other environment. We treated the data of selected environment as the test data. This way was performed on each environment and we checked the proposed method.

α was used the value of 3 states classification without environmental influence. Table V shows the result of 3 states classification with environmental influence.

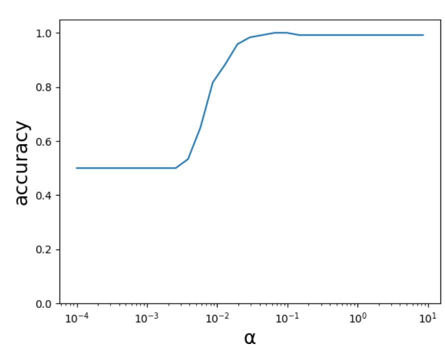

Fig. 6. The values of parameter α and the accuracy of 4 states classification in training data

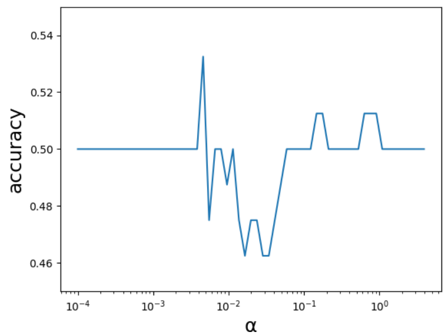

Fig. 7. The values of parameter α and the accuracy of 4 states classification in test data

TABLE V
THE ACCURACY OF 3 STATES CLASSIFICATION WITH ENVIRONMENTAL INFLUENCE

		1st	2nd	3rd	4th
subject A	No Correspondence	0.30	0.22	0.32	0.33
	Correspondence	0.35	0.37	0.43	0.40
subject B	No Correspondence	0.29	0.32	0.27	0.30
	Correspondence	0.36	0.40	0.35	0.43
subject C	No Correspondence	0.30	0.33	0.32	0.30
	Correspondence	0.41	0.43	0.41	0.38
subject D	No Correspondence	0.30	0.33	0.28	0.30
	Correspondence	0.38	0.42	0.34	0.40

$*p < 0.05$

In variance analysis, the variance ratio of all subjects fell below 5 % in this evaluation. Also, factors of classification accuracy with adapting environmental influence was significance ($p < 0.05$).

As a result, our proposed method indicated effective in 3 states classification with environmental influence.

D. 4 states classification with environmental influence

We confirmed 4 states classification with environmental influence as well as 3 states classification with environmental influence. Table VI shows the result of 4 states classification with environmental influence.

TABLE VI
THE ACCURACY OF 4 STATES CLASSIFICATION WITH ENVIRONMENTAL INFLUENCE

		1st	2nd	3rd	4th
subject A	No Correspondence	0.40	0.41	0.45	0.41
	Correspondence	0.46	0.48	0.51	0.49
subject B	No Correspondence	0.43	0.38	0.40	0.45
	Correspondence	0.51	0.45	0.50	0.53
subject C	No Correspondence	0.45	0.38	0.48	0.45
	Correspondence	0.50	0.45	0.55	0.50
subject D	No Correspondence	0.40	0.43	0.46	0.48
	Correspondence	0.50	0.45	0.51	0.52

$*p < 0.05$

In variance analysis, the variance ratio of all subjects fell below 5 % in this evaluation. Also, factors of classification accuracy with adapting environmental influence was significance ($p < 0.05$).

As a result, our proposed method indicated effective in 4 states classification with environmental influence.

V. CONCLUSION

In the present study, we proposed a classification method considering environmental influence. Our classification method is based on Bayesian inference. The method learns probability distributions. We evaluated the method. We used prepared data of multiple environments and attempted to respond to environmental influence by seeking the optimum environment.

In the experiment, we measured EEG and verified the performance of 4 state classification (rock-paper-scissors action and no action) with adapting environmental influence. Also, we performed classification under various the environment. The classification accuracy of our proposed method is higher and the method indicates significance. The method showed effectiveness to adapt environment influence.

Our proposed method will be necessary to apply to other signals such as sound data and financial data in the future. In this paper, we deal with subject physical condition and shifting measuring position as environmental changes. However, there are many environmental change (e.g. weather and battery). We need to verify that our proposed method can adapt various environmental changes.

REFERENCES

[1] T. Moriyama, H. Saito, S. Ozawa,"Evaluation of the Relation between Emotional Concepts and Emotional Parameters in Speech," *The Institute of Electronics, Information and Communication Engineers*, Vol.J82-D No.4 pp.703-711, 1999.

[2] T. Kitahara, M. Goto, H. Okuno,"Musical Instrument Identification Considering Pitch-dependent Characteristics of Timbre: A Classifier Based on F0-dependent Multivariate Normal Distribution," *IPSJ Journal*, Vol.44, No.10, pp.2448-2458. Oct. 2003.

[3] R. Osaki, M. Shimada, K. Uehara, "Extraction of Primitive Motions by Using Clustering and Segmentation of Motion-Captured Data," *The Japanese Society for Artificial Intelligence*, Vol.15.5, pp.878-886, Sep. 2000.

[4] H. Sugimura, K. Matsumoto, "Feature Pattern Extraction from Time-series Database," The 25th Annual Conference of the Japanese Society for Artificial Intelligence, Accessed 2018.4.25, Available : https://kaigi.org/jsai/webprogram/2011/pdf/507.pdf

[5] H. Tanaka, Y. Nagashima, H. Ide, "The Study of Movement Forms Discrimination by EEG Frequency Analysis," *T..IEE Japan*, Vol. 118 - C, No11, pp.1584-1589, 1998.

[6] W. Ozono, Y. Muromachi, "Development of Shopping Mode Choice Model with Bayesian Estimation Method," *Infrastructure planning review*, Vol. 25, No3, pp.717-722, Sep.2008.

[7] M. Suzuki, S. Inagaki, T. Suzuki, S. Hayakawa, N. Tsuchida, "Estimation of Switching Point in Human Driving Behavior Based on Eye-movement and Bayesian Estimation," *IIC-07-75*, pp.29-34, Mar. 2007.

[8] S. Sakai, S. Okajima, S. Izumi, A. Iwasaki, "Baysian Inference of Fatigue Life Estimated by Inspection Data," *J.Soc.Mat.Sci.,Japan*, Vol.54, No.3, pp.308-313, Mar. 2005.

[9] Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag, Berlin, Heidelberg, 2006.

[10] N. Ueda, "Bayesian Inference Algorithms - Approximation Methods for High-Dimensional Integrals," *Transactions of the Japanese Society for Artificial Intelligence*, Vol.19, No.6, pp.656-663, Nov. 2004.