Development of Machine Learning Clustering Method for Signal Processing *

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 high precision classifier for multivariate signal. The method deals with environmental influences. In this proposal technique, we express the features of the classification target and the environment in population parameters of probability distribution. We estimate the population parameters by using Bayesian Inference. Selecting similar environment properly by Bayesian Decision Rule, we try to consider environmental influences.**

In evaluation experiment, we verify that proposed method has high classification performance, and we prove that the our method adapt environmental influences from experimental results.

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

I. Introduction

We can accumulate the large data of various type because measuring technique is developing nowadays. The demand of technique 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 identifying and predicting the state. The techniques apply wide range of signal data such as finance, voice and biological.

As one example, there is a research that learns the relation between the characteristics of sound and the emotions. The research identifies feelings perceived from voice [1]. This research estimates feature of measurement voice signal using principal component analysis (PCA) and factor analysis, and the relation between audio signal and feeling is captured by the feature.

In addition, there are some applied researches. One research classifies types of the instrument from audio signal. The other research identifies 10 types of actions of radio gymnastics by using three dimensional time series data [2,3]. The research of the instrument classification uses Bagging that gets classification results from majority vote of multiple classifiers. The research of action classification uses Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). It classifies the 10 types of actions by using Bayesian Theory.

In the medical field of applied researches, the research uses the platelet data as input, and identifies patients with hepatitis B and patients with hepatitis C. Also, there is a research that classifies the action of rock-paper-scissors by Mahalanobis distance. The research uses one point Electromyogram (EEG) in measurement [4,5].

In this way, there are a lot of researches that analyze the signal by using machine learning. The method is expected that it is more expanded application fields in the near feature. Therefore, it is important to establish the technique of machine learning.

By considering diversification of data measurement, it is necessary to deal with environmental influences. In other words, accuracy of machine learning may decrease under external factors (weather and physical condition) when we obtain data of measuring instrument.

Carrying out measurement that are not affected by the environmental influences is better to improve the quality of data, but it is almost impossible because there are much restrictions on the measurement. Therefore, the robust analysis against environmental influences is necessary.

In this work, our proposed method is learning algorithm of high precision classifier. The method is based on Bayesian inference. The aim of method is to adapt environmental influences. Our proposed method classifies states expressing the features of the classification target and the environment population parameters of probability distribution. There are examples of studies using Bayesian Inference that estimated shopping mode choice, driving behavior intention using information of driver's line sight [6,7,8].

Our proposed method can deal with environmental influences to choice the optimum environment. In the experiment, we verify the performance of proposed method by using EEG data. This paper shows up a brief analysis of adapting environmental influences for the multivariate signal and evaluate such methods.

II. Proposal of state identification method

In this paper, when we obtain datasets N with D dimensions, that is acquired from measuring instruments C, we suggest the method to identify states K considering environmental influences.

A. Problem setting

X is an observed dataset N with D dimensions, that is acquired from measuring instruments C,

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

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 it has been given measuring instrument c. Also, teacher label of environment has been given by a measurement environment

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

for an environment type M. Teacher labels of states

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

represents states K of identification target. It is assumed that each $\boldsymbol{X}^{(\mathrm{c})}$ is given by

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

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

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

where $\boldsymbol{e}$ denotes the the measurement environment type and $\boldsymbol{z}$ denotes the identification target state$\boldsymbol{X}_{(\boldsymbol{e},\boldsymbol{z})}$.

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

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

In suggest method, we use Bayesian inference to learn features of each $\boldsymbol{X}_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})}$. We make an assumption the probability distribution following the multivariate Gaussian distribution, and estimate the mean vector and the variance-covariance matrix.

B. Learning rule based on Bayesian inference of probability distribution

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

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

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

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

Then, $\mathcal{N}_D(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1})$ expresses probability density function of D dimensional Gaussian distribution for mean vector μ and precision matrix $\boldsymbol{\Lambda}^{-1}$.

Also,$\mathcal{W}_D(\boldsymbol{\Lambda}|\boldsymbol{W}, \nu)$ denotes probability density function of D dimensional Wishart distribution. $\boldsymbol{I}_D$ is the identity matrix of size D.

The hyper parameters of posterior probability are defined by

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

$$\beta_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})} = \beta_0 + N_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})}, \tag{11}$$

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

$$\nu_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})} = \nu_0 + N_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})}. \tag{13}$$

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

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

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

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

where $N_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})}$ is number of $\boldsymbol{X}_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})}$. Also, $\bar{\boldsymbol{x}}_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})}$ denotes the mean vector of $\boldsymbol{x}_n^{(\mathrm{c})}$ and $\boldsymbol{S}_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})}$ denotes the the variance-covariance matrix of $\boldsymbol{x}_n^{(\mathrm{c})}$. And $(\boldsymbol{e}_n, \boldsymbol{z}_n)$ is obtained 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 the environment $\boldsymbol{e}$ and state of identification target $\boldsymbol{z}$ is given in advance, we can obtain supporting degree of measurements data $\{ \boldsymbol{x}^{(\mathrm{c})} \}$. When there is no imbalance such as the particular situation, $P\left(\boldsymbol{e}, \boldsymbol{z} | \boldsymbol{X}, \{ \boldsymbol{x}^{(\mathrm{c})} \}\right)$ is defined by

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

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

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

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

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

$$\tag{22}$$

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

$$S_D(\boldsymbol{x}|\boldsymbol{m}, \boldsymbol{Q}, f) = \frac{\Gamma(\frac{f}{2} + \frac{D}{2})}{\Gamma(\frac{f}{2})}|\boldsymbol{Q}|^{\frac{1}{2}}\left(1 + \frac{1}{f}(\boldsymbol{x} - \boldsymbol{m})^T\boldsymbol{Q}(\boldsymbol{x} - \boldsymbol{m})\right)^{-\frac{f}{2}-\frac{D}{2}} \tag{23}$$

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

$$\sum_{(\boldsymbol{e},\boldsymbol{z})\in\mathbb{T}_M\times\mathbb{T}_K} P\left(\boldsymbol{e}, \boldsymbol{z}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}\right) = 1. \tag{24}$$

If M becomes large number, the measuring environment of $\left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}$ will 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 $\mathbb{T}_M$.

There are multiple methods that classify the state from the supporting degree by eq.(18). As a simple example, we can consider the method that the classification result equals to environmental state with the maximize supporting degree.

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

$$(\boldsymbol{e}_1^*, \boldsymbol{z}_1^*) = \underset{\boldsymbol{e},\boldsymbol{z}\in\mathbb{T}_M\times\mathbb{T}_K}{\arg\max}\ P\left(\boldsymbol{e}, \boldsymbol{z}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}\right) \tag{25}$$

where $\boldsymbol{z}_1^*$ denotes the result of state classification and $\boldsymbol{e}_1^*$ denotes the 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 $\boldsymbol{e}_2^*$ based on the maximum a posteriori probability method is given by

$$\begin{aligned}\boldsymbol{e}_2^* &= \underset{\boldsymbol{e}\in\mathbb{T}_M}{\arg\max}\, P\left(\boldsymbol{e}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})}_{\mathrm{cal}} \right\}, \boldsymbol{z}_{\mathrm{cal}}\right) \\ &= \underset{\boldsymbol{e}\in\mathbb{T}_M}{\arg\max}\, P\left(\boldsymbol{e}, \boldsymbol{z}_{\mathrm{cal}}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})}_{\mathrm{cal}} \right\}\right).\end{aligned} \tag{26}$$

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

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

$$\boldsymbol{z}_2^* = \underset{\boldsymbol{z}\in\mathbb{T}_K}{\arg\max}\, P\left(\boldsymbol{e}^*, \boldsymbol{z}|\boldsymbol{X}, \left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}\right). \tag{27}$$

This identification method is expected to perform stable and highly accurate classification against changes of measuring environment because the method is not peripheralize.

In the next section, we evaluate performance of eq. (25) and eq. (27) using EEG.

III. Numerical experiment

A. Outline

In this study, we used ULTRACORTEX MARK 4 of headset and Cyton Board produced by OpenBCI. Then, we obtained EEG data and evaluate the performance of proposed method to apply the identifying state problems. This problems are total of 4 states classification including the state of no action and states of rock-paper-scissors.

As a target of the experiment, we pick up EEG data as signal. Data of multiple measuring sites can be expressed as multidimensional data. EEG may not always be able to acquire data of the same characteristics because EEG data are biological data that change the feature depend on user's physical condition and shifting measuring position. It is dangerous that change in features of data leads to lower accuracy when we predict motion by machine learning.

We suggest the robust analysis method that deal with environmental influences such as the attachment of measuring instrument and user's physical condition. Our method is based on theory of Section 2.

Also, it is difficult to acquire enormous biological data because obtaining the biological data is large burden for people. Therefore, it is necessary to analyze something from small amount of acquired data. Our proposed method can analyze small amount of observed data.

Through the experiments, we verify that the suggested method can responds to environment influences and analyze something from small amount of data. we prove that the method has the effectiveness.

B. Experimental method using EEG data

In the experiment, a subject attached the EEG and play rock-paper-scissors action following instructions of the created slide (Fig. 1). All created slide is displayed for 1 second. A subject becomes standby state when "3" and "2" of slides are displayed for 2 seconds. A slide of rock-paper-scissors is displayed for 1 second as instruction slide.

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

Measuring sites of EEG is followed ten-twenty electrode system (Fig. 2). We acquired EEG data using 8 measurement sites (Fp1, Fp2, C3, C4, P7, P8, O1, O2).

The experiment was carried out during 4 days (2018.5.9, 5.10, 5.11, 5.14) because it is considered that the environmental influences depend on physical condition of a subject and shifting measurement sites.

We define a set data for 4 seconds from the standby state to the rock-paper-scissors action state in the experiment. We obtained 100 sets of EEG data in a day. We acquired 400 sets of EEG data during 4 days.

C. Application fo proposed method to EEG data

When we analyze EEG data, there are two way. One way takes the reference point of potential from the earlobe. The other subtracts the average value of each measuring sites. In

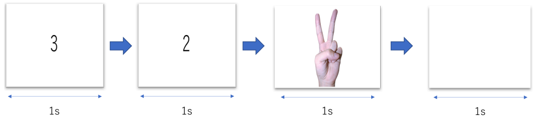

Fig. 1. Flow of creating slide

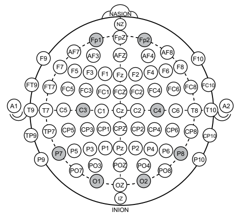

Fig. 2. ten-twenty electrode system

this study, we adopted the way subtracting the average value of each measuring sites. $\boldsymbol{S}(t) = [S^{(1)}(t), S^{(2)}(t), \cdots S^{(c)}]^{\mathrm{T}}$ is given by EEG data of measuring sites C. Then, $\boldsymbol{\mu}(t)$ is obtained by

$$\boldsymbol{\mu}(t) = \frac{1}{C}\sum_{c=1}^{C} \boldsymbol{S}^{(c)}(t). \tag{28}$$

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

$$\boldsymbol{S}_{\mathbf{0}}^{(c)}(t) = \boldsymbol{S}^{(c)}(t) - \boldsymbol{\mu}(t)\mathbf{1}. \tag{29}$$

Hamming window was applied to the acquired data. We obtained power spectra by using fast Fourier transform (FFT). The frequency band used for analysis was 3-50Hz. we divided the frequency band into 12 bands as the feature quantity for classification. The 12 dimensional feature vectors composed $\boldsymbol{S}_{\mathbf{0}}^{(c)}(t)$ of power in each band.

Also, using initial value of parameters are $\beta_0 = 0.1$, $\boldsymbol{m_0} = \mathbf{0}$, $\nu = 12$, $D = 12$, $K = 3, 4$.

D. Environmental influences

In order to check the existence of environmental influences, environmental data of each measuring sites were applied the frequency analysis. The environmental data are EEG data with the standby state. There was the difference in features of EEG data every observation day. It shows that there was an influence by the physical condition of subject and deviation from measuring points.

When we examined the presence or absence of environmental influences, a significant difference was found in the data for 4 days by graphs of frequency analysis for each measuring sites. In particular, O2 shows a clear difference. Fig. 4 shows the graph of average power spectra in the O2. This graph is composed data of 100 sets.

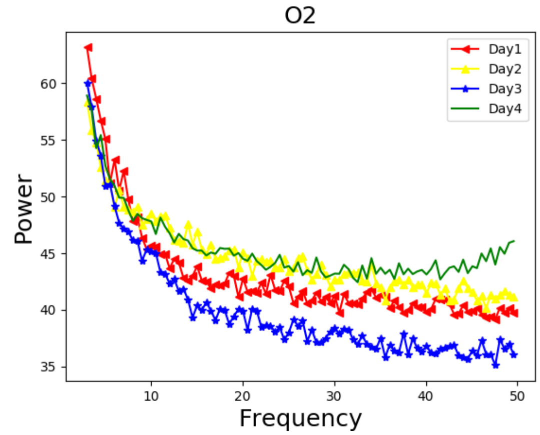

Fig. 3. Graph of frequency analysis of CH8(O1)

IV. The result of numerical experiment

To classifying the rock-paper-scissors actions, we used 1 second data while instruction slide is displayed. 2 seconds data of the standby state is 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.

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

As shown Table 1 and Table 2, the number of the training data and the test data is used the 6 types of verification. The first sets of 60 sets of acquired data was used as training data. Also, the remaining 40 sets was 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.

The 120 in Table 2 shows test data of rock-paper-scissors classes for 3days in not adapting environmental influences. Also, the 240 in Table 2 shows test data of rock-paper-scissors classes and environment class for 3days

When we execute the classification adapting environmental influences, 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 WITHOUT ENVIRONMENTAL INFLUENCES

		Not considering environmental influences
3 class	training data	60
	test data	40
4 class	training data	120
	test data	80

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

		Adapting environmental influences	Not adapting environmental influences
3 class	training data	60	60
	test data	40(the optimum environment)	120(for 3days)
4 class	training data	120	80
	test data	80(the optimum environment)	240(for 3days)

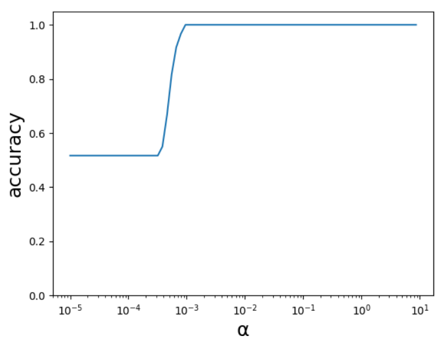

Fig. 4. The values of parameter α and the accuracy of 3 states classification in training data

A. 3 states classification without environmental influences

Table 3 shows the result of the classification without environmental influences.

TABLE III
THE ACCURACY OF 3 STATES CLASSIFICATION WITHOUT ENVIRONMENTAL INFLUENCES

	Day1	Day2	Day3	Day4
α	0.000794	0.000114	0.000154	0.000384
train(accracy)	0.81	0.51	0.66	0.55
test(accuracy)	0.38	0.53	0.43	0.41

In this study, changing parameter is only α. The values of α and the classification accuracy of the train data are shown in Fig. 6. The values of α and classification accuracy of the test data are shown in Fig. 7. Day 2 obtained the highest accuracy of 53 %.

Since the accuracy of the test data was relatively low and the classification accuracy of the train data was high, we presume that the overfitting occurred in Day1. 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 without environmental influences

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

Fig.8 shows the value of α and classification accuracy graphs of the train data. Fig.9 shows the value of α and classification accuracy graphs of the test data.Day 3 observed the highest accuracy of 55 %.

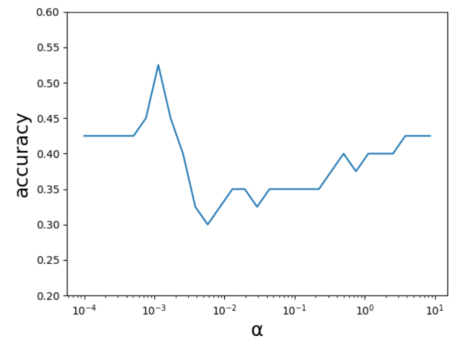

Fig. 5. The values of parameter α and the accuracy of 3 states classification in test data

TABLE IV
THE ACCURACY OF 4 STATES CLASSIFICATION WITHOUT ENVIRONMENTAL INFLUENCES

	Day1	Day2	Day3	Day4
α	0.0195	0.0046	0.0017	0.0011
train(accracy)	0.55	0.75	0.52	0.52
test(accuracy)	0.52	0.53	0.55	0.50

C. 3 states classification with environmental influences

Table 3 shows the classification results of 3 states classification with environmental influences. We verified two classifications that were not adapting environmental influences and adapting environmental influences.

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

α was used the value of 3 states classification without environmental influences. Table 5 shows the result of 3 states classification with environmental influences. The mean of classification accuracy with response to environmental influences

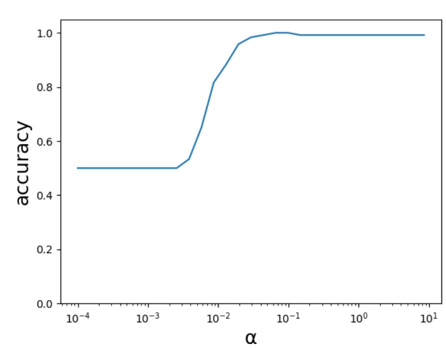

Fig. 6. The values of parameter α and the accuracy of 4 states classification in training data

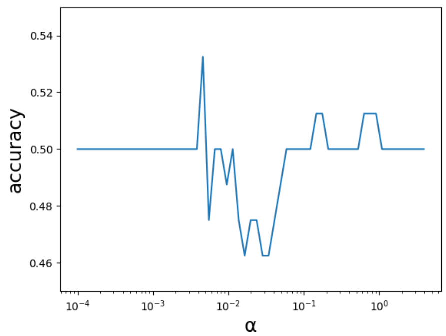

Fig. 7. The values of parameter α and the accuracy of 4 states classification in test data

TABLE V
THE ACCURACY OF 3 STATES CLASSIFICATION WITH ENVIRONMENTAL INFLUENCES

	Day1	Day2	Day3	Day4
Not adapting environmental influences	0.30	0.22	0.32	0.33
Adapting environmental influences	0.35	0.37	0.43	0.40

$*p < 0.05$

was 38.75 %. The mean of classification accuracy without response to environmental influences was 29.75 %. In variance analysis, the variance ratio became 2.2 % in this evaluation. Also, factors of classification accuracy with correspondence to environmental influences was significance ($p < 0.05$).

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

D. 4 states classification with environmental influences

We confirmed 4 states classification with environmental influences as well as 3 states classification with environmental influences. Table 6 shows the result of 4 states classification with environmental influences. The mean of classification accuracy with response to environmental influences was 48.25 %. The mean of classification accuracy without response to environmental influences was 42.75 %. In variance analysis, the variance ratio became 2.3 % in this evaluation. Also, factors of classification accuracy with correspondence to environmental influences was significance ($p < 0.05$).

TABLE VI
THE ACCURACY OF 4 STATES CLASSIFICATION WITH ENVIRONMENTAL INFLUENCES

	Day1	Day2	Day3	Day4
Not adapting environmental influences	0.40	0.41	0.49	0.41
Adapting environmental influences	0.46	0.48	0.51	0.49

$*p < 0.05$

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

V. CONCLUSION

In the present study, we proposed a classification method considering environmental influences. 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 environment influences by seeking the optimum environment.

In the experiment, we measured EEG data and verified to respond to environment influences classifying rock-paper-scissors action and no action using acquired data.

We performed classification under various the environment. The classification accuracy we proposed method is higher and the method indicate significance. Our proposed method showed effectiveness to respond to environment influences in classification in consideration of environmental change.

It is necessary to apply to other signals such as sound data and financial data in the future.

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] 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
[3] 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
[4] H. Sugimura, K. Matsumoto, "Feature Pattern Extraction from Time-series Database,"
URL: https://kaigi.org/jsai/webprogram/2011/pdf/507.pdf, Accessed 2018.4.25
[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 Behavir 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