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

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

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

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

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

In suggest method, we use 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, 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 $P_0(\boldsymbol{\mu}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}, \boldsymbol{\Lambda}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})})$ is defined by

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

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

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

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 D dimensional identity.

And hyper parameters of posterior probability is defined 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 $\boldsymbol{S}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ is 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 number of $\boldsymbol{X}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$, $\bar{\boldsymbol{x}}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ denotes the mean vector of $\boldsymbol{x}_n^{(\mathrm{c})}$, $\boldsymbol{S}^{(\mathrm{c})}_{(\boldsymbol{e},\boldsymbol{z})}$ denotes the the variance-covariance matrix of $\boldsymbol{x}_n^{(\mathrm{c})}$. And $(\boldsymbol{e}_n, \boldsymbol{z}_n)$ is given 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 known, we can obtain supporting degree of measurements data $\{ \boldsymbol{x}^{(\mathrm{c})} \}$. In the case that 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

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

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

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

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

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

(22)

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

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

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

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

In the case that M is big and learning data can be prepared under various environments, 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 is 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 we consider to the case that 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 $\left\{\boldsymbol{x}^{(\mathrm{c})}\right\}$.

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

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

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

III. Evaluation of proposed method using EEG data

A. Outline of evaluation experiment

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

EEG data includes signal as a example. Data of multiple measuring sites can be expressed as multidimensional data. EEG is not always able to acquire data of the same characteristics because EEG data is biological data that change the feature due to user's physical condition and shifting measuring position. There is a danger that change in features of data leads to lower accuracy when we predict motion by machine learning.

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

Also, it is difficult to acquire enormous biological data because the burden on people is large to acquire data. Therefore, it is necessary to analyze with small amount of getting data. Our suggest method can analyze with small amount of obsearved data.

Through evaluation experiments, we verify that the suggest method responds to environment influences and analyze with small amount of data. we indicate that the method has the effectiveness.

1) Experimental method: In the evaluation experiment, a subject attached the electroencephalograph and play rock-paper-scissors action following instructions of the created slide (Fig. 1). All slide of the created slide is displayed for 1 second. A subject become standby state when "3" and "2" of slides is displayed for 2 seconds. A slide of rock-paper-scissors slide displayed for 1 seconds as instruction slide. A subject put out the hand that will win the rock-paper-scissors hand of the instruction slide when white slide is displayed after instruction slide.

Measuring sites of EEG followed ten-twenty electrode system (Fig. 2). We acquired EEG using 8 measurement sites (Fp1, Fp2, C3, C4, P7, P8, O1, O2). Also, Fig. 3 shows the state of the experiment.

The number of subject is one person. Evaluation experiment were carried out over 4 days (2018.5.9, 5.10, 5.11, 5.14) because it is considered that there are environmental influences such as physical condition of a subject and shifting measurement sites.

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

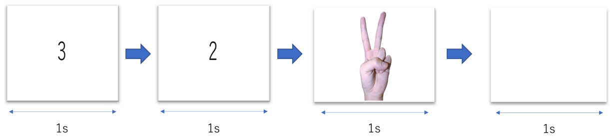

Fig. 1. Flow of creating slide

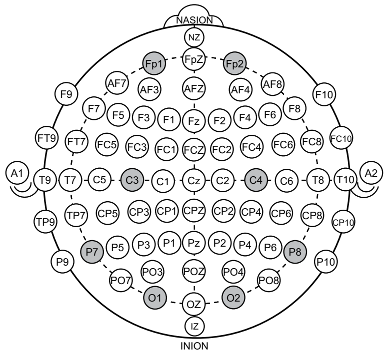

Fig. 2. ten-twenty electrode system

2) Environmental change: In order to check the presence or absence of environmental influences, Environmental data of each measuring sites were applied the frequency analysis. Environmental data is the standby state of EEG data. The fact that there was the difference in features of EEG data were observed every observation day. It suggests that there was an influence by the physical condition of subject and attachment or detachment of the measuring instrument.

When examining 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, CH3 (C3) and CH8 (O2) showed a clear difference. Fig.4 and Fig.5 show the average graphs of CH3(C3)'s and CH8(O2)'s power spectra for 100 sets.

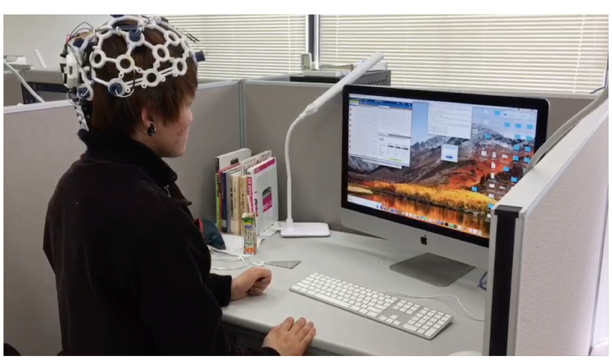

Fig. 3. Wearing of an electroencephalograph

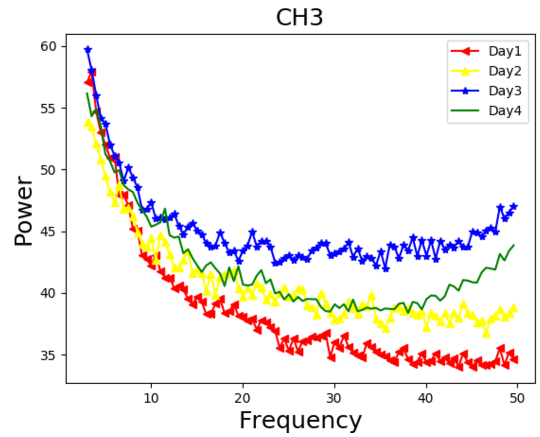

Fig. 4. Graph of frequency analysis of CH3(C3)

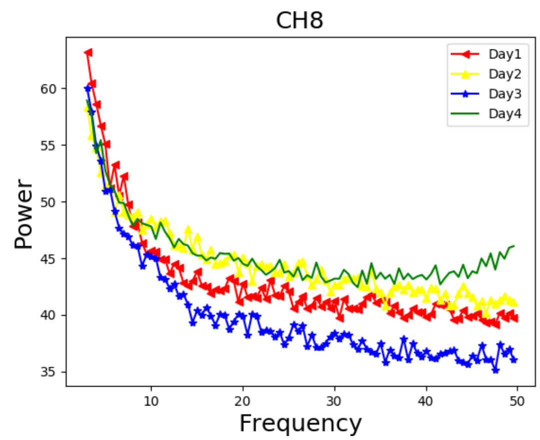

Fig. 5. Graph of frequency analysis of CH8(O1)

B. Application of proposal identification method to EEG data

When analyzing EEG, there are two case that take the reference point of potential from the earlobe and subtract the average value of each measuring sites. In this study, we adopted the case subtracting the average value of each measuring sites.

$\boldsymbol{S}(t) = [S^{(1)}(t), S^{(2)}(t), \cdots S^{(c)}]^{\mathrm{T}}$ is given from EEG data of measuring sites C. Then, $\boldsymbol{\mu}(t)$ is given by

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

In each measurement sites, obtained EEG data $\boldsymbol{S}_{\boldsymbol{0}}^{(c)}(t)$ to subtract signal average value $\boldsymbol{\mu}(t)$

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

Hamming window was applied to the acquired data. We obtained power spectra 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. Used 12 dimensional feature vector composed $\boldsymbol{S}_{\mathbf{0}}^{(c)}(t)$ of power in each band.

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

C. Performance result of the proposed method

To classifying the rock-paper-scissors actions, we used 1 second data while instruction slide is displayed. 2 seconds data of standby state is used as the environment data. We carried out two kinds of verification. One verification is 3 classification of rock-paper-scissors. The other verification is 4 classification of rock-paper-scissors and no action. In addition, we also two kinds of verification that the case whether environmental influences can be ignored in each that verification.

In the case that environmental influences can be ignored, there is not measuring instrument with attachment or detachment. The acquired data of first sets of 60 sets was used as training data on the same day. The remaining 40 sets was used as test data. We obtained classification accuracy of each day by using the train data and the test data.

In the case that environmental influences can not be ignored, evaluation was performed by using the state classification of the environment.

1) 3 states classification without environmental influences: Table 1 shows the result of classification for 4 days of 2018.5.9(Day1),2018.5.10(Day2), 2018.5.11(Day3),2018.5.14(Day4)

TABLE I
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 is shown Fig. 6. The values of α and classification accuracy of the test data is shown Fig. 7. Day 2 got 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, it could be presumed the overfitting 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.

2) 4 states classification without environmental influences: In the 4 states classification, it included the no action class in addition to rock-paper-scissors classes. we verified whether EEG data is changing by setting the no action class when a subject was playing rock-paper-scissors. Table 2 shows the result of 4 states classification.

Fig.8 shows the value of α and classification accuracy graphs of the train data. Fig.9 shows the value of α and

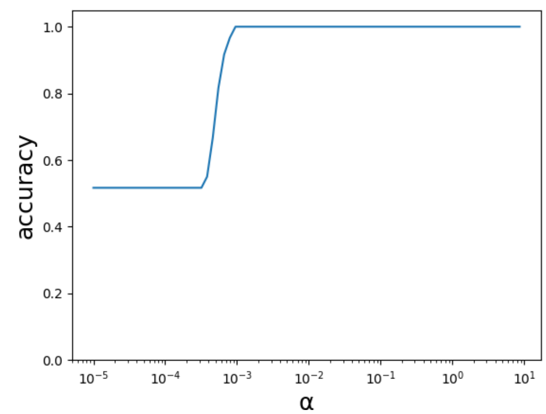

Fig. 6. The values of parameter α and the accuracy of 3 states classification in train data

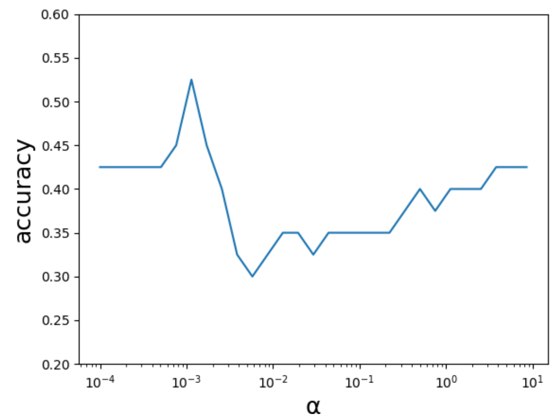

Fig. 7. The values of parameter α and the accuracy of 3 states classification in test data

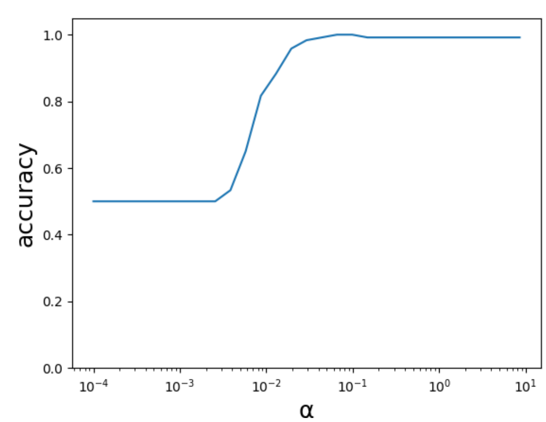

Fig. 8. The values of parameter α and the accuracy of 4 states classification in train data

TABLE II
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

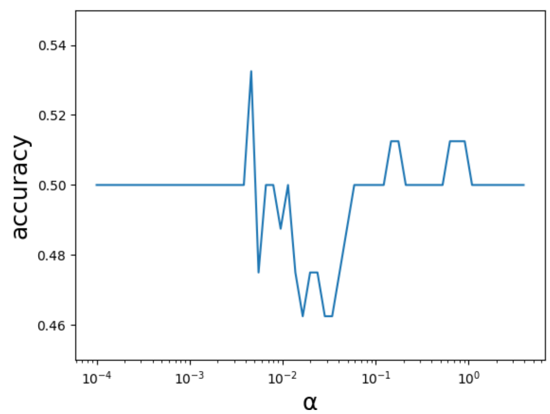

Fig. 9. The values of parameter α and the accuracy of 4 states classification in test data

classification accuracy graphs of the test data.Day 3 observed the highest accuracy of 55 %.

3) Environmental states determination: In this study, acquired data of no action for 4 days were used as calibration data. We estimated the optimum environment based on eq. (26) against that calibration data. 100 datasets of each day was used as the calibration data. The result of the estimation of the optimum environment is shown in Fig. 10.

We estimated the optimum environment of each day's data like the result that the optimum environment of Day 1 is Day 3. Our classification was carried out to estimate optimum

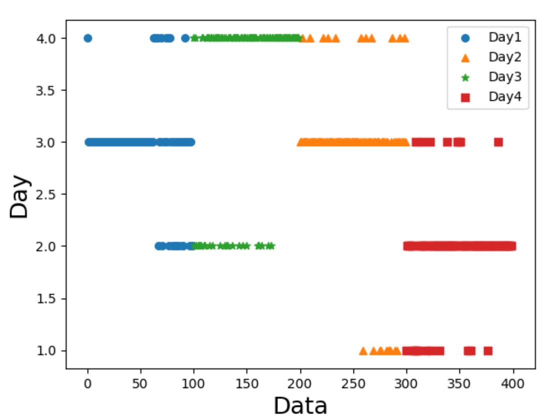

Fig. 10. Environmental condition determination

environment.

4) 3 states classification with environmental influences: Table 3 shows the classification results of 3 states Classification with environmental influences. In this case, there were two classification that were not adapting environmental influences and adapting environmental influences. Also, α was used the value of 3 states Classification without environmental influences.

TABLE III
THE ACCURACY OF 3 STATES CLASSIFICATION WITH ENVIRONMENTAL INFLUENCES

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

$*p < 0.05$

The average of classification accuracy with response to environmental influences was 38.75 %. The average 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.

5) 4 states classification with environmental influences: We obtained classification accuracy in case of responding to 4 classification with environmental influences. This 4 states classification includes no action class. Also, we classified after choosing the optimum environment like 3 states classification with environmental influences.

TABLE IV
THE ACCURACY OF 4 STATES CLASSIFICATION WITH ENVIRONMENTAL INFLUENCES

	Day1	Day2	Day3	Day4
test(No response to environmental change)	0.40	0.41	0.49	0.41
test(Responding to environmental change)	0.46	0.48	0.51	0.49

$*p < 0.05$

The average of classification accuracy with response to environmental influences was 48.25 %. The average 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 ($p < 0.05$). As a result, Our proposed method indicated effective in 4 states classification with environmental influences

IV. CONCLUSION

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

In evaluation experiment, we measured EEG data and verified to respond to environment influences classifying rock-paper-scissors action using acquired data. In the case of 3 states classification and 4 states classification, 4 days EEG data was divided into training data and test data to obtain the accuracy of the proposed method.

We classified under various the environment. In classifying under the different, we classified 3 states and 4 states after selected the optimum environment by environmental condition determination. The accuracy of classification average were 38.75 % in 3 states classification and 48.5 % in 4 states classification, respectively. When we compared the case that not considering environmental influences, classified 3 states and 4 states after selected the optimum environment by environmental condition determination. The accuracy of classification average were 29.25 % in 3 class classification and 42.75 % in 4 class classification. 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, 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, Oct.2003

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

[5] H. Tanak, Y. Nagashima, H. Ide, "The Study of Movement Forms Discrimination by EEG Frequency Analysis,"T..IEE Japan, Vol. 118 -C, No11, 98

[6] W. Ozono, Y. Muromachi, "Development of Shopping Mode Choice Model with Bayesian Estimation Method," Infrastructure planning review, Vol. 25, No3, 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, Mar.2005