Development of Machine Learning Clustering Method for Signal Processing *

*Note: Sub-titles are not captured in Xplore and should not be used

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*—The accuracy of machine learning may remarkable decrease under condition that there are external factors (weather and physical condition) as well as the instrument calibration and getting only data influenced by its variation. In this paper, we suggest learning algorithm of high precision classifier for multivariate signal that deal with environment influences. In this proposal technique, we assume the estimation of the probability distribution an axis from a small number of data by using Bayesian Inference and express the features of the discrimination target and the environment based on the population parameters of probability distribution and estimate its population parameters**

Also, we try to deal with environment influences by selecting properly similar environment using getting dataset of Bayes Decision Rule and using its learning result In evaluation experiment, we validate to get high classification performance using EEG by suggestion method and prove that suggestion method corresponds to influence the environment to compare the case of not considering environment influences.

***Index Terms*—Signal Processing, Bayes Inference, EEG**

I. Introduction

We can accumulate the multifarious of large data because measuring technique is developing nowadays. Also, The demand of technique that analyze comprehensively acquired data and application example of machine learning what is known as automatic analysis technique by the computer is increasing.

In particular, Many researches of machine learning increase to exploit wide range of signal data such as finance, voice, biometric data for purpose of identifying and predicting the state. As one example, there is a research that learn the relation between the characteristics of sound and the emotions included in voice and identify feelings perceived from voice [1]. This research presume feature basis of measurement voice signal using principal component analysis(PCA) and factor analysis and the relation between audio signal and feeling is captured by feature basis using multiple regression analysis.

In addition, there are applied researches that determine types of the instrument from audio signal and identify 10 types of actions of radio gymnastics using Three dimensional time series coordinate data [2,3]. The former uses Bagging that get classification results from output of multiple classifiers by majority vote. The latter uses method for reducing dimension of Principal Component Analysis (PCA) and Liner Discriminant Analysis (LDA) and classifier on Bayes Theory.

In medical field, Input is the platelet data and using decision tree that include in one of the machine learning classifies patients with hepatitis B and patients with hepatitis C. Also, there is a research what classify the action of rock-paper-scissors by Electromyogram(EEG) in a single point of measurement based on Mahalanobis distance [4,5].

In this way, there are a lot of case that analyze the signal using machine learning and it is important for the method of machine learning that is expected to be more expanded application field in the near future to establish.

On the other hand, Considering diversification of data measurement, it is necessary to deal with change in measurement environment and its condition. In other words, accuracy of machine learning may remarkable decrease under condition that there are external factors (weather and physical condition) as well as the instrument calibration and getting only data influenced by its variation.

Performing measurements that are not affected by the environment is better the method to improve the quality of data but it is necessary to make a robust analysis against environmental influences when there are much constraints on the measurement.

In this work, we suggest learning algorithm of high precision classifier for multivariate signal that deal with environment influences. In this proposal technique, we assume the estimation of the probability distribution an axis from a small number of data by using Bayesian Inference and express the features of the discrimination target and the environment based on the population parameters of probability distribution and estimate its population parameters.

Also, we try to deal with environment influences by selecting properly similar environment using getting dataset of Bayes Decision Rule and using its learning result. we validate to get high classification performance using EEG by suggestion method and prove that suggestion method corresponds

Identify applicable funding agency here. If none, delete this.

to influence the environment to compare the case of not considering environment influences.

II. Proposed state identification method

In this paper, when we obtain set of N data are acquired from measuring instruments C each capable of measuring D dimensional data, we suggest the method to identify states K by considering environmental influences by using them.

A. Problem setting

We obtain variables X that is observed set of N data are acquired from measuring instruments C each capable of measuring D dimensional data.

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

where we have defined

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

where $\boldsymbol{X}^{(c)}$ is data that have been given measuring instrument c. Also, teacher label have been giving measurement environment types M along with it.

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

Teacher labels representing identification target states K

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

It is assumed that each is given

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

where we have defined

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

where $\boldsymbol{X}_{(\boldsymbol{e},\boldsymbol{z})}$ is a set in which data corresponding to the measurement environment type $\boldsymbol{e}$ and the identification target state $\boldsymbol{z}$ is extracted

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

where each $\boldsymbol{X}_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})}$ is the data set with features according to its 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 suppose the probability distribution that the data follows is the multivariate Gaussian distribution and estimate mean vector and variance-covariance matrix that is its parameters.

B. Learning rule based on Bayesian inference of probability distribution

Under the assumption of the preceding paragraph, where we have defined

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

where $\boldsymbol{\mu}_{(\boldsymbol{e},\boldsymbol{z})}$, $\boldsymbol{\Lambda}_{(\boldsymbol{e},\boldsymbol{z})}$ has been mean vector and precision matrix for the data set $\boldsymbol{X}_{(\boldsymbol{e},\boldsymbol{z})}$.

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

where $Q(\boldsymbol{\mu}_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})}, \boldsymbol{\Lambda}_{(\boldsymbol{e},\boldsymbol{z})}^{(\mathrm{c})} | \boldsymbol{X}^{(\mathrm{c})})$ have been the posterior distribution.

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

Also,$\mathcal{W}_D(\boldsymbol{\Lambda}|\boldsymbol{W}, \nu)$ have been Probability density function of D dimensional Wishart distribution and $\boldsymbol{I}_D$ is D dimensional identity.

where we have defined

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

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

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

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

Each has been parameters that defines the posterior distribution. where

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

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

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

Then,

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

When the environment $\boldsymbol{e}$ and Identification target state $\boldsymbol{z}$ has been known, we indicate degree of supporting measurements data $\left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}$. In the case that there is no imbalance such as when the particular situation is likely to occur, given 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 $\boldsymbol{m}^*, \boldsymbol{Q}^*, f^*$ have been given by

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

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

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

$$\tag{22}$$

Also, S_D have been multivariate t-distribution of degree of freedom f, given by

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

$\Gamma(\cdot)$ is Gamma function. Then,

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

where $P\left(\boldsymbol{e},\boldsymbol{z}|\boldsymbol{X},\left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}\right)$ have been standardized.

In the case that M is big and learning data can be prepared under various environments, the measurement environment of $\left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}$ will be found the similar environment within measurement environment type of learning data. Therefore, the measurement environment $\boldsymbol{e}$ of $\left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}$ have been treated as the element of $\mathbb{T}_M$.

There are multiple methods methods that classify the state from the degree of support by (18) to be considered. As a simple example, it is conceivable to adopt a method of determining the state of the state of the environment that maximizes the degree of support obtained as the discrimination result. Therefore,

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

$\boldsymbol{z}_1^*$ is the result of state determination.

Next, considering to the case that measurement data $\left\{ \boldsymbol{x}^{(\mathrm{c})}_{\mathrm{cal}} \right\}$ and identification target state $\boldsymbol{z}_{\mathrm{cal}}$ have been known for calibration. we can identify the measurement environment 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 by calibration data, the state of the continuous measurement data $\left\{ \boldsymbol{x}^{(\mathrm{c})} \right\}$ have been classified.

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

This identification method have been expected to perform stable and highly accurate classification against changes in properties of data due to measurement environment because that method have not done peripheralize due to adding identification procedure of the similar environment compared (25).

In the next section, performance evaluation of (25) and (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 got EEG and evaluated the performance of suggest method to apply the problems identifying states that total of 4 states including the state of no action and state of rock-paper-scissors.

EEG includes signal as a example and measurement data of a plurality of measurement sites can be expressed as multidimensional data. EEG is not always able to acquire data of the same characteristics each time because EEG is biological data that change the feature due to user's physical condition and measurement position shifts somewhat when detaching. There is a danger that change in features of data leads to lower accuracy when predicting motion by machine learning using EEG.

Therefore, we suggest robust analysis method that deal with environment influences such as attachment of measuring instrument and user's physical condition based 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 getting data.

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

1) Experimental method: In the evaluation experiment, a subject have the electroencephalograph attached and perform rock-paper-scissors action in accordance with instructions on the created slide(Fig.1) and analyze getting data. All slide of the created slide is displayed for 1 second. A subject become

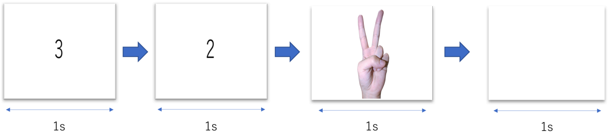

Fig. 1. Flow of creating slide

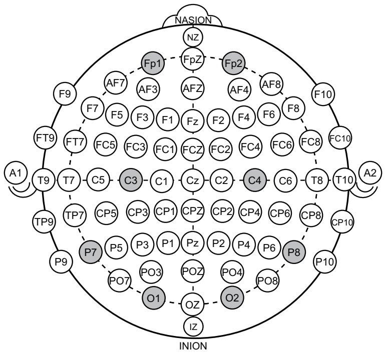

Fig. 2. ten-twenty electrode system

standby state when ”3” and ”2” of slides is displayed for 2 seconds and a slide of rock-paper-scissors slide displayed for 1 seconds as instruction slide. A subject put out the hand of the prize that will win the rock-paper-scissors hand of the instruction slide when white slide is displayed after instruction slide is displayed.

Measurement sites of EEG followed ten-twenty electrode system (Fig.2) and 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 conducted 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 the subject and shifting measurement sites due to attachment/detachment of the measurement instrument.

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

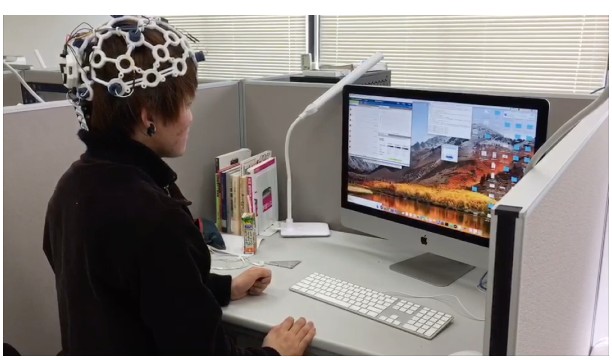

Fig. 3. Wearing of an electroencephalograph

2) Environmental change: In order to check the presence or absence of environmental influences, frequency analysis was performed for each measurement sites for environmental data. Environmental data is the standby state of EEG data. The fact that differences in features of such EEG data were observed every observation day suggests that there was an influence due to the physical condition of subject and slight deviation due to attachment and detachment of the measurement instrument.

When examining the presence or absence of environmental influences based on the data difference on each measurement dates, a significant difference was found in the data for 4 days by graphs of frequency analysis for each measurement 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.

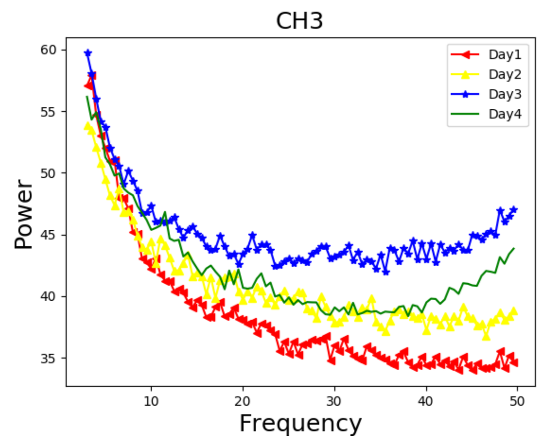

Fig. 4. Graph of frequency analysis of CH3(C3)

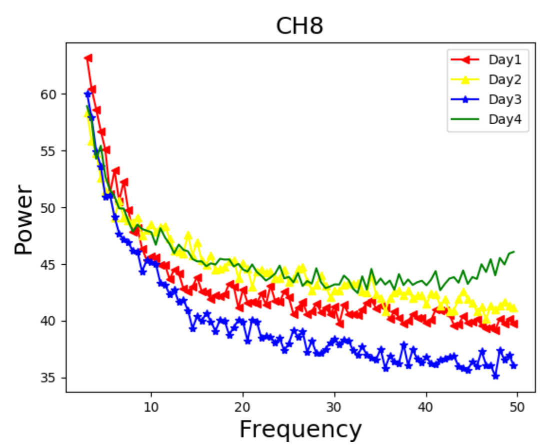

Fig. 5. Graph of frequency analysis of CH8(O1)

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 measurement sites. In this study, we adopted the latter.

In EEG data of measurement site$C$, $\boldsymbol{S}(t) = [S^{(1)}(t), S^{(2)}(t), \cdots S^{(c)}]^{\mathrm{T}}$ is given. 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, getting EEG data $\boldsymbol{S}_{\mathbf{0}}^{(c)}(t)$ to subtract signal average value $\boldsymbol{\mu}(t)$

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

Hamming window was applied to the acquired data and got 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 and 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

The data used to classify the rock-paper-scissors action was 1 second data while instruction slide is displayed and 2 seconds data of standby state is used as the environment data of no action. we made two kinds of verification that 3 classification of rock-paper-scissors and 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 attachment / detachment and acquired data on the same measurement day first set of 60 sets of data for one day was used as training data and the remaining 40 sets were used as test data to obtain classification accuracy of each day.

On the other hand, evaluation was performed using the state classification of the environment of the suggest method. we trained with data of one day and data of the other day as test data and selected the environment based on optimal environment identification.

1) Classification without responding to 3 class environmental changes: Table 1 shows the result of classification for acquired data for 4 days of 2018.5.9(Day1),2018.5.10(Day2), 2018.5.11(Day3),2018.5.14(Day4)

TABLE I
CLASSIFICATION ACCURACY IN THE CASE WHERE ENVIRONMENTAL INFLUENCES CAN BE IGNORED IN 3 CLASSIFICATION

	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

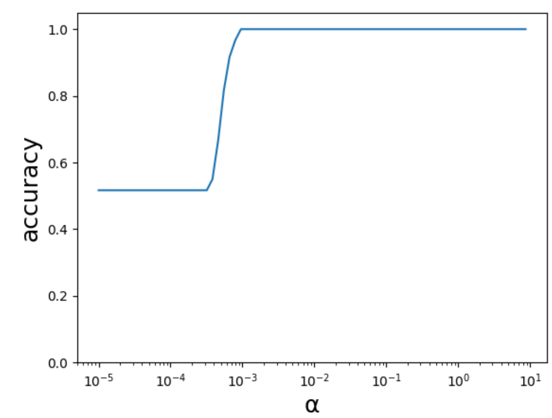

Fig. 6. Parameter α of the training data of 3 class classification and graph of accuracy

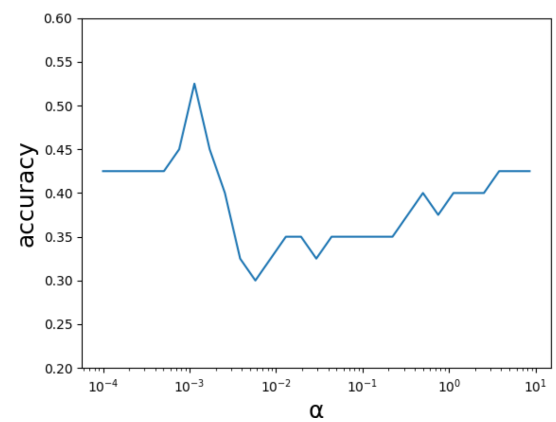

Fig. 7. Parameter α of the test data of 3 class classification and graph of accuracy

In this study, changing parameter is only α. The values of α and the classification accuracy of the training data and the classification accuracy of the test data are shown in Fig.6, 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 training data is high, it could be seen that overfitting occurs and it could not be classified successfully. On the other hand, because the accuracy of the test data of the other three days and the classification accuracy of the training data did not change, it could be considered that classification could be done without causing overfitting.

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

Fig.8 and Fig.9 show the value of α and classification accuracy graphs of the training data and the classification

TABLE II
CLASSIFICATION ACCURACY WITHOUT RESPONSE TO ENVIRONMENTAL CHANGE

	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	0.5

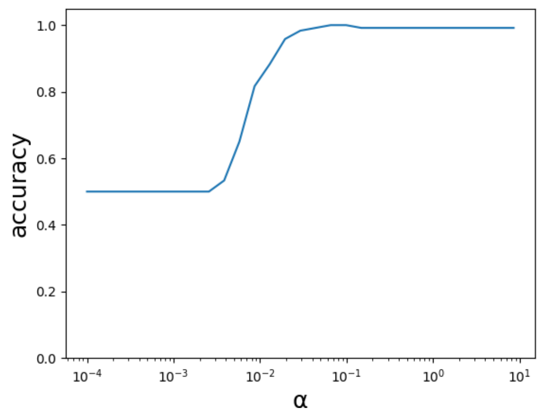

Fig. 8. Parameter α of the training data of 4 class classification and graph of accuracy

accuracy of the test data. Day 3 got the highest accuracy of 55 %

3) Environmental condition determination: In this study, acquired data of no action for 4 days were used as calibration data. We estimated the optimum environment based on (26) against that calibration data. Data of 100 sets of data on 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 for Day 1 is

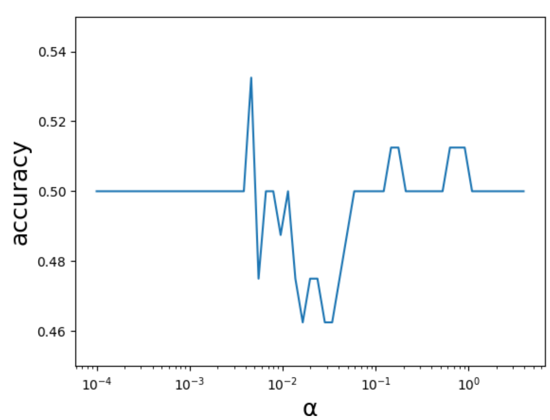

Fig. 9. Parameter α of the test data of 4 class classification and graph of accuracy

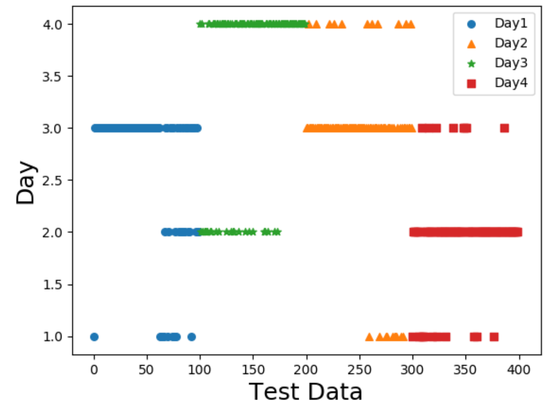

Fig. 10. Environmental condition determination

Day 3. Therefore, when dealing with environmental influences, classification was carried out using data of the estimated optimum environment.

4) Classification of correspondence to three classes of environmental change: Table 3 shows the classification results when classifying data using 3 days data other than the optimum environment as test data. Also, the value of α was used when the environmental influences could be ignored in 3 classification.

TABLE III
CLASSIFICATION ACCURACY OF CORRESPONDENCE TO ENVIRONMENTAL CHANGE

	Day1	Day2	Day3	Day4
test(No response to environmental change)	0.30	0.22	0.32	0.33
test(Responding to environmental change)	0.35	0.37	0.43	0.4

$*p < 0.05$

As a result of dispersion analysis, the dispersion ratio became 2.2 % in this evaluation. Also, Factors of classification accuracy with correspondence to environmental change and correspondence to environmental change were significant ($p < 0.05$).

As a result, Since the average of classification accuracy with response to environmental change was 38.75 % and the average of classification accuracy without response to environmental change was 29.75 %, the classification that corresponded to environmental change was more accurate.

5) Classification of correspondence to 4 classes of environmental change: We sought classification accuracy in case of responding to 4 classification of environmental influences including standby state. Also, we classified after choosing the environment by identifying the optimum environment as in the case of responding to three class environmental changes.

As a result of dispersion analysis, the dispersion ratio became 2.3 % in this evaluation. Also, Factors of classification accuracy with correspondence to environmental change

TABLE IV
CLASSIFICATION ACCURACY OF CORRESPONDENCE TO ENVIRONMENTAL CHANGE

	Day1	Day2	Day3	Day4
test(No response to environmental change)	0.4	0.41	0.49	0.41
test(Responding to environmental change)	0.46	0.48	0.51	0.49

$*p < 0.05$

and correspondence to environmental change were significant ($p < 0.05$).

As a result, Since the average of classification accuracy with response to environmental change was 48.25 % and the average of classification accuracy without response to environmental change was 42.75 %, the classification that corresponded to environmental change was more accurate.

IV. CONCLUSION

In this paper, We proposed a classification method of signals considering environmental influences. Our classification used learning probability distribution based on Bayes Inference and prepared training data of multiple environments and attempted to respond to environment changes by seeking the optimum environment by Bayes Theory.

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

We classified under the different environment and the same environment. In classifying under the different, we classified 3 class and 4class after selected the optimum environment by environmental condition determination. The result of classification's average accuracy was 38.75 % in 3 class classification and 48.5 % in 4 class classification. When we compared the case that not considering environmental influences, classified 3 class and 4class after selected the optimum environment by environmental condition determination. The result of classification's average accuracy was 29.25 % in 3 class classification and 42.75 % in 4 class classification. The classification accuracy that coped with each environmental change was higher and it was able to show significance. Our suggest method showed effectiveness to respond to environment influences in classification in consideration of environmental change. We also showed that classification can be done with a small amount of data

We verified the proposed method using EEG data in this study but it is necessary to verify whether it can be applied to other signals such as sound data and financial data.

REFERENCES

[1] Tsuyoshi MORIYAMA, Hideo SAITO, Shinji 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 https://kaigi.org/jsai/webprogram/2011/pdf/507.pdf, 2018.4.25

[2] Ryura Osaki，Mituomi Shimada, Kuniaki Uehara，"Extraction of Primitive Motions by Using Clustering and Segmentation of Motion-Captured Data"，The Japanese Society for Artificial Intelligence, Vol.15.5(2000 年 9 月)

[3] Tetsuro Kitahara, Masataka Goto, Hiroshi 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] Hiroshi SUGIMURA，Kazunori MATSUMOTO，"Feature Pattern Extraction from Time-series Database", https://kaigi.org/jsai/webprogram/2011/pdf/507.pdf, 2018.4.25

[5] Tanaka Hisaya, Nagashima Yuji, Ide Hideto, "The Study of Movement Forms Discrimination by EEG Frequency Analysis" ,T..IEE Japan, Vol. 118 - C, No11, 98