ONLINE BRAIN-MACHINE INTERFACE WITH AUTOMATIC DETERMINATION OF STOPPING TIME OF TRAINING PHASE

Yumi Dobashi,* Atsushi Takemoto,** Shu Shigezumi,* Takumi Shiraki,* Katsuki Nakamura,** and Takashi Matsumoto*

Abstract

Electroencephalography (EEG) signals are one of the most popular signals used for brain–machine interfaces (BMIs). EEG-based BMI methods often work in batch mode, where a user must conduct the learning phase for a pre-determined period of time. This paper proposes an EEG-based sequential BMI system in which (i) the machine can determine when to end the learning phase automatically by monitoring the learning progress using the sequential error rate (SER) as an evaluation index and (ii) sequential learning in both the brain and the machine in a cooperative manner is employed. In the proposed approach, called brain–machine co-learning, subjects learn how to use the system by means of real-time visual feedback, whereas the machine learns the subjects' EEG signals by Bayesian sequential learning. The SER refers to the average classification error rate windowed over a short time period, which was proposed in Hara *et al.*, Sequential error rate evaluation of SSVEP classification Problem with Bayesian sequential learning, *Proc. 10th IEEE Int. Conf. on Information Technology and Applications in Biomedicine*, Corfu, Greece, November 2–5, 2010, and it represents the status of Bayesian sequential learning in real time. In our proposed approach, subjects can use the system while eliminating unnecessary training. The proposed system was tested against a steady-state visual-evoked potential classification problem. The training phase varied for each subject and was sometimes short, yet satisfactory, leading to high classification accuracy.

Key Words

Biomedical signal processing, biomedical computing, brain–machine interface, electroencephalography

* Department of Electrical Engineering and Bioscience, Graduate School of Science and Engineering, Waseda University, Tokyo 169-8555, Japan; e-mail: {shigezumi10, shiraki10}@ matsumoto.eb.waseda.ac.jp, bandaid23@gmail.com, takashi@ matsumoto.elec.waseda.ac.jp
** The Primate Research Institute Kyoto University, Inuyama, Aichi 484-8506, Japan; e-mail: takemoto.atsushi.3r@kyoto-u. ac.jp, katsuki@pri.kyoto-u.ac.jp

1. Introduction

Electroencephalography (EEG) monitors electrical brain activity observed on the scalp by electrodes. EEG-based brain–machine interfaces (BMIs) are one of the most studied methods among the many promising non-invasive BMIs that have been proposed. This is mainly due to their fine temporal resolution, ease of use, and relatively low setup cost. However, each BMI method naturally has disadvantages. One disadvantage with an EEG-based BMI is its susceptibility to noise, which has motivated a variety of machine learning algorithms for decoding EEG signals, and there have been significant advances in the area.

There are several ways to categorize BMIs, one being to categorize them as operating in either the batch mode or the sequential (online) mode. In the batch mode, the EEG data collected from a subject is divided into two subsets: training data and test data. The former is used for training the machine learning algorithm, whereas the latter is used to evaluate the algorithm's capability to predict the subject's intention [1]–[3]. There are several aspects of batch-mode learning that need to be improved:

- First, it is non-trivial how much data should be used for training and how much data should be preserved for testing. Also, the number of necessary training data items may depend on each subject.
- Second, with the batch-mode learning, by definition, one cannot perform sequential evaluations of the subject's predictive performance as time evolves.
- Third, the batch-mode learning presumes that the data is stationary, *i.e.*, that the subject's physical condition and/or the environment around the subject does not change over the period of the experiment.
- Fourth, there is no key for a subject indicating his or her performance in real time. This often causes difficulty in using BMIs.

By using the sequential error rate (SER), which was proposed in [4], this paper proposes two paradigms in Bayesian sequential learning.

- Automatic determination of when to stop the training phase: For this purpose, the SER is monitored in real time.
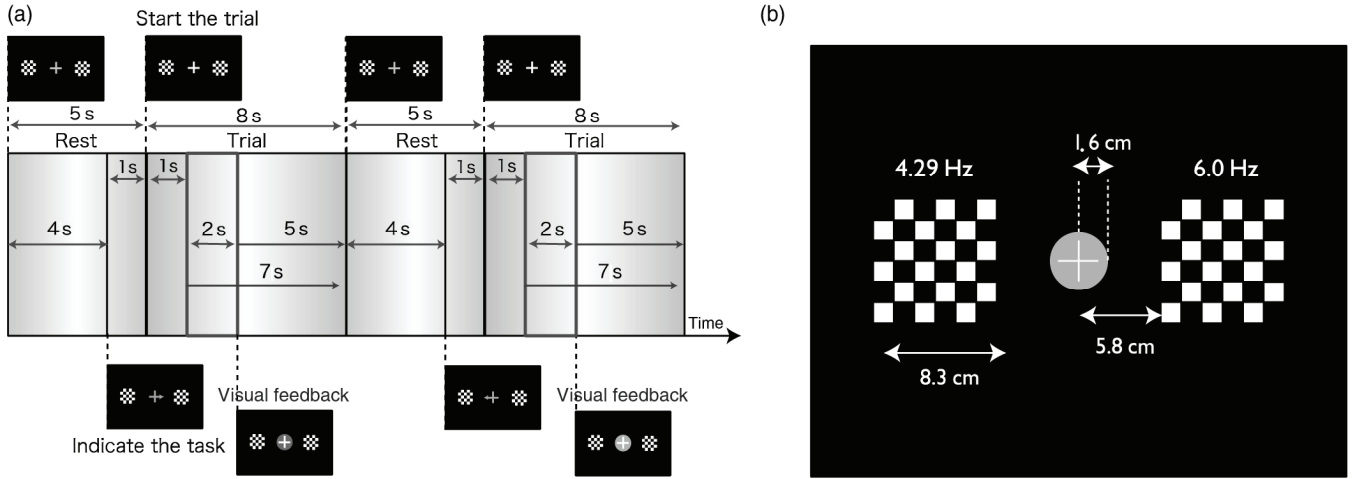
Figure 1. (a) Trial flow of the proposed system showing display timing and (b) monitor display with two flickering checkerboard stimuli.

- Brain–machine co-learning: Learning in both the brain and the machine in a cooperative manner. In the training phase, subjects learn how to use the BMI by means of real-time visual feedback, whereas the machine learns the subjects' EEG signals via a sequential Bayesian learning algorithm.

It should be noted that the sequential learning in [4] is performed each time trial data arrives, where trial data consists of 3.0 s of EEG data. In the latter, learning is performed 10 times within a trial in real time, which enables visual feedback in the proposed BMI system. We tested the proposed system against the steady-state visual-evoked potential (SSVEP) classification problem. One of the main reasons for using SSVEP is that it is relatively stable in the presence of several possible EEG signals. SSVEP-based BMI has been a well-known technology since the 1990s [5].

2. Related Work

For the batch-mode learning, we cite the survey paper reported in [1] instead of citing individual papers. Allison *et al.* [6] report the appropriateness of an SSVEP-based EEG system for those subjects with no experience in a very noisy field setting. Martinez *et al.* [3] propose an SSVEP-based online BMI system with visual neurofeedback, but besides the different approach for preprocessing the EEG data, its algorithm is that of a fuzzy neural network classifier, namely, adaptive neuro fuzzy inference system (ANFIS), not Bayesian sequential learning. Brown *et al.* [7] is a paper on a linguistics problem using a Bayesian framework. There are also reports on Bayesian sequential learning for EEG signals in [4] and [8], where the sequential Monte Carlo (SMC) is used for implementation. In [4], the SER was proposed for evaluating sequential learning.

Recently, some studies have focused on changes in the subject's condition during the use of brain-computer interface (BCI). Satti *et al.* [9] use a covariate shift minimization to alleviate the non-stationary nature of EEG data. Chung *et al.* [10] propose an EEG-based BCI system that

takes into account the subject's tiredness and uncertainty of output during use, and displays a menu with two options, "guide" and "exit", during the test phase to ask the subject whether to do more training. This approach can adapt to the subject's physical condition and/or the environment around the subject, but it depends on the subject's own decision. Our approach can automatically evaluate the certainty of the system in the training phase.

3. Subjects and System

3.1 Subjects

Five volunteers (aged 19–23, four males and one female) participated in the study. Written informed consent was obtained from each subject on forms approved by the ethical committee of Waseda University. All subjects were healthy, with no past history of psychiatric or neurological disorders.

3.2 Experimental Setup

EEG data were recorded with two active electrodes (O1,O2) according to the international 10–10 system and referenced to the left earlobe with a digitization rate of 500 Hz. Each subject was seated in a comfortable chair 60 cm in front of a monitor. The flow of task events is shown in Fig. 1(a), and the stimulus is illustrated in Fig. 1(b).

The EEG data was measured with a Polymate AP1124, a multi-purpose portable bio-amplifier recording device, manufactured by TEAC Corporation, Tokyo, Japan. The device is equipped with 24 channels with a maximum sampling frequency of 1 kHz. In addition to electroencephalograms, eyeball movement and other external signals can be measured. The dimensions are W 90 mm × H 44 mm × D 158 mm, and the weight is 300 g. The device is powered by a battery.
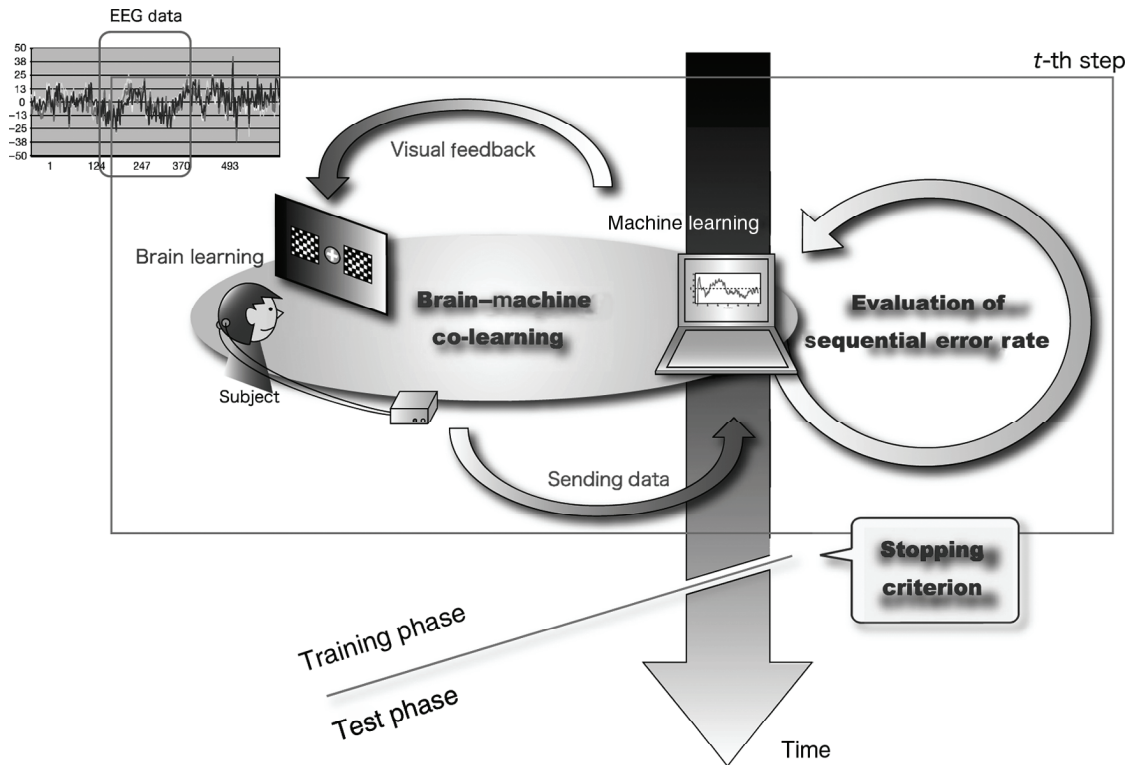
Figure 2. The scheme of the proposed BMI system. The system involves brain–machine co-learning with visual feedback, and the SER is evaluated every time an item of data arrives. To automatically determine when to stop the training phase, the SER is constantly monitored. The training phase is stopped when the SER is below a threshold for a certain period (the stopping criterion), and then the system switches to the test phase automatically.

3.3 Procedure

Each subject completed five sessions. One session consisted of two parts, a training phase and a test phase. In the training phase, the subject underwent up to 20 trials. Two flickering checkerboard stimuli were presented on the monitor at frequencies of 4.29 Hz on the left and 6.0 Hz on the right. The subject usually fixated on a central fixation cross. When an arrow replaced the cross, the subject was instructed to move his or her eyes to the checkerboard indicated by the arrow for 8.0 s. During this time, 3.0 s after the subject moved his/her eyes to the checkerboard, a coloured circle appeared in the centre of the monitor for visual feedback. The colour of the circle changed every 0.5 s according to the likelihood described in Section 5.4. The colour had 10 gradations. If the likelihood was high, the colour became bluish, whereas if the likelihood was low, the colour became reddish. This colour change was used by subjects to learn how to use the system and to get an indication of their EEG condition in real time. The direction of the arrow was alternated between left and right in sequence. After 8.0 s, the subject returned his/her eyes to the central fixation cross and rested for 5.0 s. This sequence constituted one trial. The colour of the central fixation cross was green when the subject should rest, and changed to white while the subject moved his/her eyes to the checkerboard.

When the prediction error was below a certain threshold in two consecutive trials (details are described in the next section), causing the system to automatically stop the training phase, a honey-coloured circle was displayed in the centre of the monitor during the rest time, so that the subject would know when to start the test phase.

The test phase was used to evaluate the machine's capability to predict the subject's intention. During this time, the subject performed the same task as in the training phase, but the machine could no longer learn the subject's EEG signal. The direction of the arrow was selected at random in this phase.

4. System Implementation

4.1 Flow of the Trial

Figure 2 shows the implementation of our system. As explained in Section 3.2, EEG data was acquired from two channels (O1,O2) located in the occipital eye field. One trial lasted for 8.0 s. Data in the first 0.0–1.0 s was deleted to eliminate the influence of eyeball movements on the EEG. The EEG data was converted into a frequency spectrum by a discrete Fourier transform (DFT) with a Hanning window. From each frequency spectrum, the two frequency components (4.29 and 6.0 Hz) were selected, so that the feature vector was four-dimensional. The feature vector was sent to the SMC algorithm every 0.5 s for 5.0 s after the DFT time window of 2.0 s. In one trial, there were 10 steps. In each step, the past 20 items of observation data were sent to the SMC.

Each time EEG data arrived in 0.5 s, the likelihood defined in Section 5.4 was computed and was converted into one of ten different colours for display at the centre of the monitor. This served as the real-time visual feedback to the subject.

Note that the sequential learning process was not reset even when moving to the next trial.

4.2 Sequential Error Rate

The SER was calculated in every step. The SER of the t-th data item in this study is defined by:

$$\text{SER} := \frac{1}{M} \sum_{t'=t-M+1}^{t} I(y_{t'} \neq y_{t',pred}) \qquad (1)$$

where y_t ($y_{t'}$) is the true class, and $y_{t,pred}$ ($y_{t',pred}$) is the predicted class defined by (8). The function $I(A)$ represents an indicator function defined as 1 when A is *true* and 0 when A is *false*. In our experiment, $M = 20$. We averaged the error over the interval $\{1, 2, \ldots, t\}$, where $t < 20$.

4.3 Criterion for Stopping Training Phase

From the start of the experiment, the SER was calculated in each step for automatically evaluating when to stop the training phase. A decreasing SER means that the brain–machine co-learning is functioning well, whereas an increasing SER means that the brain–machine co-learning is not working. Stopping the training phase when the SER is decreasing is effective for ensuring that the BMI performs prediction correctly. In this paper, we set a parameter μ as the threshold for estimating when to stop the training phase. Within one trial containing 10 datasets, if the SER of the latter half, that is, 6 out of 10 datasets, is below μ, the system considers that the brain–machine co-learning is functioning well, in at least the first trial. When this event occurs a second time in a row, the system stops the training phase automatically.

5. Algorithm

Let $\boldsymbol{x}_t \in \boldsymbol{R}^d$ be the feature vector at the t-th data item, where d represents the dimension of $\boldsymbol{x}_t$, which, in our paper, is the DFT spectrum of a single trial EEG. Let $y_t \in \{0, 1\}$ be the binary class label of each trial, where 0 corresponds to the right flickering image, and 1 corresponds to the left flickering image. Our purpose here is to learn parameters associated with the basis function, to be defined shortly, and to predict the subject's intention given SSVEP data, each time data arrives.

5.1 Basis Function and Classifier

This study utilizes the nonlinear basis function used in Ref. [3] which we denote by $f(\boldsymbol{x}_t, \boldsymbol{w}_t)$ where $\boldsymbol{w}_t$ denotes the parameter vector associated with the basis function. To associate the quantity of $f(\boldsymbol{x}_t, \boldsymbol{w}_t)$ with the class label, consider:

$$P(y_t | \boldsymbol{x}_t, \boldsymbol{w}_t) := Be(y_t; \Phi(f(\boldsymbol{x}_t; \boldsymbol{w}_t))) \qquad (2)$$

where Φ is a function that monotonically maps the real numbers onto [0,1]. Of the several possible choices of Φ, we choose $\Phi(u) := \frac{1}{1 + \exp(-u)}$, and consider:

$$P(y_t | \boldsymbol{x}_t, \boldsymbol{w}_t) := (\Phi(f(\boldsymbol{x}_t; \boldsymbol{w}_t)))^{y_t} (1 - \Phi(f(\boldsymbol{x}_t; \boldsymbol{w}_t)))^{(1-y_t)} \qquad (3)$$

5.2 Parameter Search Stochastic Dynamics

We perform a sequential stochastic search of the parameter $\boldsymbol{w}_t$ each time trial data is acquired:

$$P(\boldsymbol{w}_t | \boldsymbol{w}_{t-1}, \gamma) := \frac{1}{Z_w(\gamma)} \exp\left(-\frac{\gamma \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|^2}{2}\right) \qquad (4)$$

where Z_w represents the normalization constant. This amounts to searching for a new value $\boldsymbol{w_t}$ based on the previous value $\boldsymbol{w_{t-1}}$, but in a random walk manner:

$$\boldsymbol{w_t} = \boldsymbol{w_{t-1}} + \boldsymbol{\mu_t} \qquad (5)$$

where $\boldsymbol{\mu_t}$ is a Gaussian random variable with zero mean and variance $1/\gamma$. We note that γ determines the parameter search region for $\boldsymbol{w_t}$.

5.3 Posterior Distributions

Letting $\boldsymbol{x}_{1:t} := (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t)$, $y_{1:t} := (y_1, \ldots, y_t)$, one can derive its sequential posterior distribution at step t:

$$P(\boldsymbol{w}_t | \boldsymbol{x}_{1:t}, y_{1:t}) = \frac{P(y_t | \boldsymbol{x}_t, \boldsymbol{w}_t) P(\boldsymbol{w}_t | \boldsymbol{x}_{1:t-1}, y_{1:t-1})}{\int P(y_t | \boldsymbol{x}_t, \boldsymbol{w}_t) P(\boldsymbol{w}_t | \boldsymbol{x}_{1:t-1}, y_{1:t-1}) d\boldsymbol{w}_t} \qquad (6)$$

The second factor in the numerator is given by:

$$\begin{aligned} &P(\boldsymbol{w}_t | \boldsymbol{x}_{1:t-1}, y_{1:t-1}) \\ &= \int P(\boldsymbol{w}_t | \boldsymbol{w}_{t-1}) P(\boldsymbol{w}_{t-1} | \boldsymbol{x}_{1:t-1}, y_{1:t-1}) d\boldsymbol{w}_{t-1} \end{aligned} \qquad (7)$$

At the $t + 1$-th step, let the EEG data $\boldsymbol{x}_{t+1}$ be given. Then the prediction at this step amounts to computing the predictive probability:

$$\begin{aligned} &P(y_{t+1} | \boldsymbol{x}_{1:t+1}, y_{1:t}) \\ &= \int P(y_{t+1} | \boldsymbol{x}_{t+1}, \boldsymbol{w}_{t+1}) P(\boldsymbol{w}_{t+1} | \boldsymbol{x}_{1:t}, y_{1:t}) d\boldsymbol{w}_{t+1} \end{aligned} \qquad (8)$$

This study uses the SMC to evaluate (6) and (8) in an attempt to evaluate the SER.

5.4 Robustness Issue

Note that training data with the same class label for consecutive steps often leads to inappropriate sequential posterior samples, because appropriate posterior samples are difficult to find when only one class label is available for a whole interval. To overcome this difficulty, we regard the EEG signal for the past 20 items as the observation data
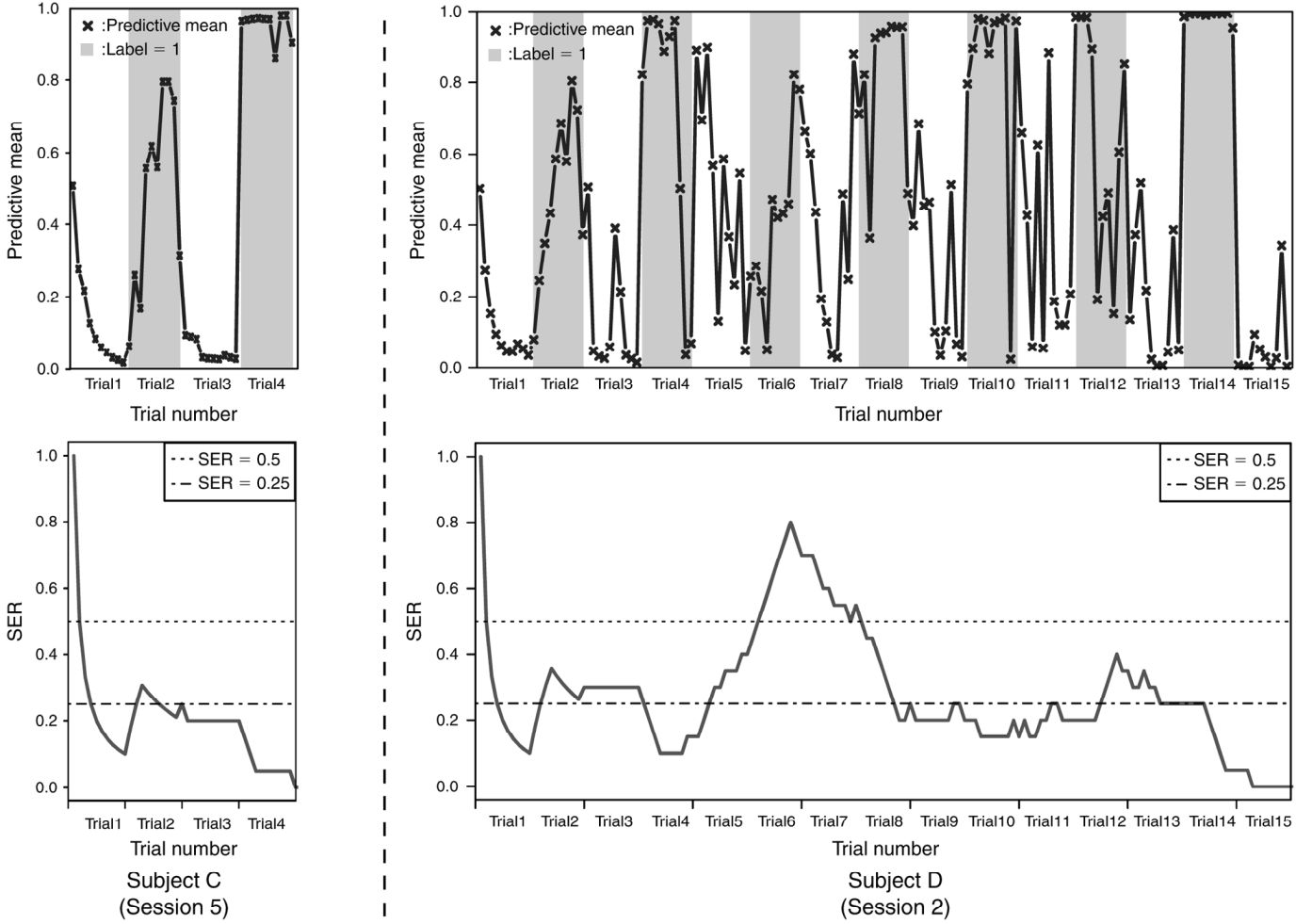
Figure 3. Prediction results of subjects C (left) and D (right). The upper figure is the trajectory of the label predictive probability. The red rectangles in the upper graph stand for the true labels when the true class labels are 1. Intervals without such rectangles indicate that the true labels are 0. The bottom graph shows the SER associated with the above data. Note that there are 10 data in a trial. The two subfigures on the right show the prediction results for Subject D, where it took the subject 15 trials to satisfy the stopping criterion, which was the longest among the subjects.

at time t, where both labels are present in these 20 items. We redefine the equation for the likelihood as follows:

$$(t > 20) \; P(y_t | \boldsymbol{x}_t, \boldsymbol{w}_t)$$

$$:= \prod_{\tau = t-19}^{t} \left(\Phi(f(\boldsymbol{x}_\tau ; \boldsymbol{w}_\tau)) \right)^{y_\tau} (1 - \Phi(f(\boldsymbol{x}_\tau ; \boldsymbol{w}_\tau)))^{(1-y_\tau)} \quad (9)$$

6. Results

This section reports the results of automatic evaluation of the SER and the degree of accuracy achieved with the proposed system. Two channels (O1,O2) in the occipital area were selected for the experiments. The time-series data from each trial was converted into a frequency spectrum by a DFT with a Hanning window of 2.0 s. From each frequency spectrum, the two frequency components (4.29 and 6.0 Hz) were selected, so that the feature vector was four-dimensional.

6.1 Experimental Settings

Various experimental settings were made, as follows. The number of samples, n, was 1,000, and the hyper-parameter,

γ, was 100. The SER threshold, μ, used to automatically estimate when to stop the training phase was set to 0.25. This value has roots in an empirical rule; for example, our past research showed that for over 80% of training data, the SER was under 0.25 in some subjects who resulted in high classification accuracy. In the case where the SER did not fall below the threshold μ, the training phase was stopped automatically. One trial lasted for 8.0 s, and rests between trials lasted for 5.0 s. One session included a training phase (up to 20 trials) and a test phase (10 trials), and so the maximum experimental time for one session was 6.5 min.

In addition to evaluating the SER, we also evaluated the classification accuracy during 10 trials of the test phase.

6.2 Experimental Results

Five subjects (A, B, C, D, and E) participated in the experiment. The EEG data of the trial was subjected to a DFT to obtain our feature vector.

The two subfigures on the left in Fig. 3 show the predictions associated with Subject C, who made a successful completion within four trials in terms of the stopping

Table 1

The prediction performance of the subjects A–E in their test phases, consisting of five sessions

		Session Number				
		1	2	3	4	5
A	Accuracy (%)	85.7	85.7	100	75.0	80.0
	#Trials for completion	8	7	7	9	10
	#Inconclusive trials	3	3	0	2	0
B	Accuracy (%)	100.0	100.0	80.0	70.0	100.0
	#Trials for completion	4	5	5	8	7
	#Inconclusive trials	0	0	0	0	0
C	Accuracy (%)	100.0	100.0	100.0	100.0	100.0
	#Trials for completion	5	3	6	3	4
	#Inconclusive trials	1	0	2	1	0
D	Accuracy (%)	100.0	87.5	100.0	100.0	100.0
	#Trials for completion	10	15	3	7	8
	#Inconclusive trials	1	2	0	0	2
E	Accuracy (%)	100.0	100.0	100.0	100.0	100.0
	#Trials for completion	12	7	5	5	5
	#Inconclusive trials	1	0	0	0	0

The data consists of accuracy (%) and the number of trials needed to satisfy the stopping criterion described in Section 4.3. The inconclusive trials indicate the trials that could not be judged as correct or incorrect.

criterion described in Section 4.3. The blue graph in the upper left of the figure is the label predictive probability, defined by:

$$z_t := \int P(y_t|\mathbf{x}_t, \mathbf{w}_t)P(\mathbf{w}_t|\mathbf{x}_{1:t-1}, y_{1:t-1})d\mathbf{w}_t$$
$$= P(y_t|x_{1:t}, y_{1:t-1}) \qquad (10)$$

The red rectangles in the bottom-left graph stand for the true labels when the true class labels are 1. Intervals without such rectangles indicate that the true labels are 0.

Table 1 summarizes the prediction performance of the subjects A–E in their test phases. The data consists of accuracy (%) and the number of trials needed to satisfy the stopping criterion. We note that the number of trials each subject must perform depended on the performance. The minimum number of trials before stopping was 3, whereas the maximum number of trials before completion of the training phase was 15. Each subject underwent 10 trials in each test phase. In one trial, 10 datasets were sent to the system and 10 label predictive probabilities (8) were received. We assumed that each trial indicated at the left was "correct" if six out of ten label predictive probabilities exceeded 0.5. Similarly, each trial indicated at the right was correct if six out of ten label predictive probabilities were less than 0.5. If five out of ten label predictive probabilities were correct and the rest were incorrect, we regarded the trial as inconclusive and excluded it when evaluating the

accuracy. We evaluated the accuracy of our proposed system by the following method, where C is the number of correct trials, and IC is the number of incorrect trials:

$$\text{accuracy} := \frac{C}{(C + IC)} \qquad (11)$$

7. Discussion

Figure 3 indicates that our proposed sequential learning algorithm, together with the SER, is functional. As shown in Table 1, the accuracies for the five subjects were reasonably high.

The left graph in Fig. 3 shows the result of one of the sessions for Subject C in the training phase. The trajectory of label predictive probability is shown at the top, and the SER is shown at the bottom. In the former, the red rectangles represent the true label when the true class labels are 1, whereas the intervals without such rectangles indicate that the true labels are 0. During trials 1–2, the label predictive probability gradually approached its true label. Note that the class labels during this period remained the same, as alluded to in Section 5.4. After this period, the label predictive probability immediately started following its true labels, which appeared to indicate that the proposed brain–machine co-learning is functional. The bottom-left graph shows the SER of Subject C in the same session as above. The SER started decreasing immediately

from the start of trial 1 until the trial ended. During trial 2, however, the SER gradually increased because another label came into the learning data. After this, the SER gradually decreased with some fluctuations. Near the end of trial 4, the SER dropped to almost 0.1. The proposed system detected this and stopped the training phase automatically. The right part of Fig. 3 shows the corresponding results for Subject D, who needed the longest training phase. In the graph of the label predictive probability shown in the upper right, sometimes the label predictive probability did not follow the true label, as observed in trials 5–7. The SER at the bottom right shows a tendency different from that of Subject C. The SER started decreasing immediately after the experiment started, like Subject C. In this case, however, even though the SER dropped to around 0.3, it did not decrease any more during trial 3. The SER then dropped to 0.1 and increased gradually. As described in Section 4.3, the system does not stop the training phase unless the SER falls below the threshold for a certain period of time for the second trial in a row. Therefore, the training phase continued. From the start of trial 14, the SER was under 0.25 for a while. The system detected this and stopped the training phase.

The number of trials of the training phase in Table 1 was between 3 and 15. From this result, we observed that the time variation of the SER could vary from session to session, which means that even for the same subject, his/her condition differed in each session. The classification accuracies of these subjects were generally high: two out of five subjects achieved 100% in all sessions, and one subject achieved the minimum number for the training phase.

This indicates the possibility that, if the subject can learn how to use the system quickly, he or she can start to use it immediately, which could be important in the practical use of BMIs.

8. Conclusion

We constructed a BMI system that automatically determines when to stop the training phase by employing a brain–machine co-learning approach involving Bayesian sequential learning and real-time visual feedback. The system was evaluated in an SSVEP two-class classification problem, where five subjects participated in the experiment. The number of trials for the training phase varied among subjects. The minimum was 3, and the maximum was 15. The accuracy of the classification was high, and two subjects achieved 100% in all sessions.

Acknowledgements

References

[1] A. Bashashati, M. Fatourechi, R.K. Ward, and G.E. Birch, A survey of signal processing algorithms in brain–machine interfaces based on electrical brain signals, *Journal of Neural Engineering*, *4*(2), 2007, R32–R57.

[2] G.R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, Steady-state visual evoked potential (SSVEP)-based communication: Impact of harmonic frequency components, *Journal of Neural Engineering*, *2*, 2005, 123–130.

[3] P. Martinez, H. Bakardjian, and A. Cichocki, Fully online multicommand brain–computer interface with visual neurofeedback using SSVEP paradigm, *Computational Intelligence and Neuroscience*, *2007*, 2007, 1–9.

[4] H. Hara, A. Takemoto, Y. Dobashi, K. Nakamura, and T. Matsumoto, Sequential error rate evaluation of SSVEP classification problem with Bayesian sequential learning, *Proc. 10th IEEE Int. Conf. on Information Technology and Applications in Biomedicine*, Corfu, Greece, November 2–5 2010.

[5] D. Zhu, J. Bieger, G.G. Molina, and R.M. Aarts, A survey of stimulation methods used in SSVEP-based BCIs, *Computational Intelligence and Neuroscience*, *2010*, 2010, 12 pages. doi:10.1155/2010/702357

[6] B. Allison, T. Lüth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Gräser, BCI demographics: How many (and what kinds of) people can use an SSVEP BCI?, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *18*(2), 2010, 107–115.

[7] P.F. Brown, V.J.D. Pietra, S.A.D. Pietra, and R.L. Mercer, The mathematics of statistical machine translation: Parameter estimation, *Computational Linguistics*, *19*(2), 1993, 263–311.

[8] J.W. Yoon, S.J. Roberts, M. Dyson, and J.Q. Gan, Adaptive classification for brain computer interface systems using sequential Monte Carlo sampling, *Neural Networks*, *22*, 2009, 1286–1294.

[9] A. Satti, C. Guan, D. Coyle, and G. Prasad, A covariate shift minimization method to alleviate non-stationarity effects for an adaptive brain–computer interface, *Pattern Recognition (ICPR)*, *20th International Conference*, Istanbul, Turkey, 23–26 August 2010, 105–108.

[10] M. Chung, W. Cheung, R. Scherer, and R.P.N. Rao, A hierarchical architecture for adaptive brain–computer interfacing, *Proc. 22nd Int. Joint Conf. on Artificial Intelligence – Vol. Two*, Barcelona, Spain, 2011, 1647–1652.

Biographies

Yumi Dobashi has received a master's degree in engineering at the department of electrical engineering and bioscience, Waseda University, Tokyo, Japan. She has been doing research on SSVEP-based BMI with Bayesian sequential learning. She has been particularly interested in constructing algorithms for online brain–machine co-learning, where the brain as well as the machine learn together in a cooperative and online manner. Implementations of the algorithms are by SMC methods.

Atsushi Takemoto received his M.A. degree in psychology from Kyoto University, Japan in 1991. He worked as an assistant professor in the faculty of integrated human studies at Kyoto University from 1994 until he joined the department of information science at the Shizuoka Institute of Science and Technology in 1998 as a lecturer. From 2004 to 2009, he worked as a researcher at

the National Institute of Neuroscience. Now, he is a researcher at the Primate Research Institute, Kyoto University. His areas of research interest are visual psychophysics, EEG, and animal cognition.

Shu Shigezumi received his B.E. degree in electrical engineering and bioscience from Waseda University, Tokyo, Japan and is currently pursuing his M.S. degree. His research interest is in Bayesian sequential learning for EEG-based BMI. In particular, he has been constructing sequential learning/prediction algorithms for SSVEP overt/covert attention experiments using SMC methods.

Takumi Shiraki received his B.E. degree in electrical engineering and bioscience from Waseda University, Tokyo, Japan and is currently pursuing his M.S. degree. His research interest is in Bayesian sequential learning for EEG-based BMI. In particular, he has been constructing adaptive methods for SSVEP overt/covert attention experiments.

Katsuki Nakamura obtained his Ph.D. degree in neurophysiology from Kyoto University in 1992. After he worked at the Primate Research Institute, Kyoto University as an assistant professor and associate professor, he joined the National Institute of Neuroscience, NCNP, where he was a director of a department. He is now a professor of a section of cognitive neuroscience, Primate Research Institute, Kyoto University. He has been interested in neural mechanisms underlying emotion, communication, memory, *etc*. His subjects are humans, macaque monkeys, and common marmosets. His methods are neurophysiology, brain imaging, experimental psychology, *etc*.

Takashi Matsumoto received his B.E. in electrical engineering from Waseda University, Tokyo, Japan, M.S. in applied mathematics from Harvard University, Cambridge, Massachusetts, and Ph.D. in electrical engineering from Waseda University, Tokyo, Japan. Currently, he is a professor at the Graduate School of Advanced Science and Engineering. His interests are in Bayesian learning algorithms for real world problems with Monte Carlo implementations. The target applications include EEG as well as NIRS based brain signal decoding, hyperspectral imaging based individual authentication and image segmentation, among others. He also attempts to construct algorithms with nonparametric prior distributions for Bayesian learning and applications. He was a recipient of the 1994 Best Paper Award of the Japanese Neural Network Society. He is a life fellow of the IEEE. Dr. Matsumoto held visiting positions at Cambridge University, UK and UC Berkeley.